

DUAL-FREQUENCY ACOUSTIC SEABED
CLASSIFICATION ON THE SCOTIAN SHELF, CANADA

ANDREW CUFF

**Dual-frequency Acoustic Seabed Classification on the Scotian Shelf,
Canada**

by

© Andrew Cuff

A thesis submitted to the
School of Graduate Studies
in partial fulfillment of the
requirements for the degree of

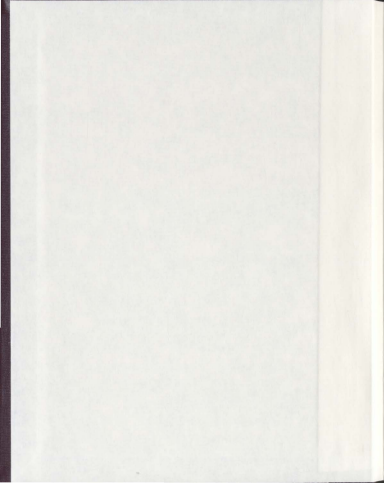
Masters of Science

Department of Geography
Faculty of Arts
Memorial University of Newfoundland

February 2011

St. John's

Newfoundland and Labrador



Abstract

Acoustic seabed classification is typically done using single acoustic frequencies, although it is known that different frequencies can potentially provide additional information on the seabed. This thesis explores the potential of using two acoustic frequencies, 38 kHz and 120 kHz, from a single beam echo sounder for acoustic seabed classification on the Scotian Shelf, Canada. The main goals were to assess the characteristics of near-nadir acoustic backscatter as a function of seabed substrate properties and acoustic frequency, analyze differences in frequency response at different spatial scales, and to compare single frequency and dual-frequency classifications of seabed substrate types. Univariate and multivariate statistical methods were used to determine the statistical characteristics of single and dual-frequency near-nadir acoustic backscatter and to identify frequency differences. Both supervised and unsupervised classification techniques were used to classify the data. Results revealed trends in dual-frequency near-nadir backscatter such as higher frequencies and larger sediments generally produce higher backscatter and more heterogeneous and rough seabeds produce variable backscatter. It was found that differences in frequency responses of backscatter occurred at scales of hundreds of meters. Finally, it was concluded that dual-frequency classification improved upon single frequency classification at near-nadir incidence angles.

Acknowledgements

I would first of all like to thank the guidance, knowledge, and dedication that my two co-supervisors, Dr. John Anderson and Dr. Rodolphe Devillers, provided me with during the research writing of this thesis. I would also like to acknowledge the commitment and input of Dr. Sam Bentley, who was also part of my committee. Gratitude goes out to Memorial University of Newfoundland, Fisheries and Oceans Canada, MITACS Accelerate, Natural Sciences and Engineering Research Council of Canada, for financial support throughout this research. I would also like to thank Chris Lang for his guidance and knowledge for helping me understand the physics behind underwater acoustics and his support during the data processing stage. Appreciation goes out to Denise Holloway and Neil Otterhead who also provided technical assistance during the data processing stage. Both Dr. Evan Edinger and Dr. Alvin Simms deserve recognition for some guidance with the statistical analysis. I would like to thank the faculty, staff, and fellow students of Memorial University's Geography Department for providing a healthy and positive work environment. Finally, I would like to thank my friends and family for their support throughout my time as a graduate student.

Table of Contents

Abstract	ii
Acknowledgements	iii
Table of Contents	iv
List of Figures	vi
List of Tables	x
List of Abbreviations and Symbols	xiii
List of Appendixes	xiv
1. Introduction	1
1.1 Context	1
1.2 Research Problem	7
1.3 Research Hypothesis and Research Questions	9
1.4 Project Objectives	12
1.5 Methods	12
1.6 Thesis Layout	14
2. Literature Review	15
2.1 Acoustic Seabed Classification (ASC)	15
2.2 Acoustic Seabed Backscatter	18
2.3 Single Beam Echo Sounders	23
2.4 Acoustic Frequency	27
3. Methods	30
3.1 Study Area	32
3.2 Data Collection	40
3.3 Data Processing	42
3.4 Univariate Acoustic Backscatter Analysis	55
3.5 Univariate Acoustic Backscatter Classification	57
3.6 Frequency Differences	58
3.7 Multivariate Acoustic Backscatter Analysis	60
3.8 Multivariate Classification	62
3.8.1 Unsupervised k-means Clustering	63
3.8.2 Supervised Discriminant Function Analysis	65

4. Results	69
4.1 Univariate Acoustic Backscatter Analysis	69
4.2 Univariate Acoustic Backscatter Classification	80
4.3 Frequency Differences	84
4.3.1 Frequency Differences Based on Classified Backscatter Values	85
4.3.2 Frequency Differences Based on Continuous Backscatter Values	94
4.4 Multivariate Acoustic Backscatter Analysis	100
4.5 Multivariate Classification	114
4.5.1 Supervised Discriminant Function Analysis	114
4.5.2 Unsupervised k-means Clustering	130
4.6 Summary of Results	134
5. Discussion	136
5.1 Univariate Acoustic Backscatter Analysis	136
5.2 Univariate Acoustic Backscatter Classification	140
5.3 Frequency Differences	141
5.4 Multivariate Acoustic Backscatter Analysis	143
5.5 Multivariate Classification	144
5.6 General Considerations	146
6. Conclusions	150
References	159
Appendix A: Interpreted Sediments Units (ISUs)	165
Appendix B: Mann-Whitney U Test Results	166
Appendix C: Quantile Classification Maps	170
Appendix D: Frequency Differences Based on Classified Backscatter Values	173
Appendix E: Frequency Differences Based on Continuous Backscatter Values	177
Appendix F: Multivariate Analysis	185
Appendix G: Balanced Backscatter Data	187

List of Figures

Figure 1: Knowledge pyramid representing levels of understanding from data to knowledge. Adapted from Poore and Chrisman (2006).	2
Figure 2: Typical acoustic signal returns from a smooth surface (interface 1) and a rough surface (interface 2). These graphs illustrate the influence of roughness on the shape of the returned acoustic signal (source: Sternlicht and de Moustier, 2003).	20
Figure 3: Acoustic backscatter response showing both the surface (i.e. interface) and volume scattering of a returned acoustic signal as a function of time (source: Sternlicht and de Moustier, 2003).	22
Figure 4: Ensonified area of a single acoustic beam as a function of time from a circle (left side) to an annulus (right side) on a flat horizontal seabed (source: Kloser, 2007).	25
Figure 5: Acoustic energy return of a SBES footprint and how it can be used to measure the near-nadir and off-nadir zones (adapted from: Anderson, 2007).	26
Figure 6: Low and high acoustic frequency waveforms (source: National Oceanic and Atmospheric Administration, NOAA).	27
Figure 7: Detailed flowchart of the research method. Bold text indicates previous work not done by the author.	31
Figure 8: Study area displaying Western Bank (rectangle) on the outer Scotian Shelf. Banks are represented by areas shallower than 91 m depth (50 fathoms). Shaded contours represent depths in meters.	33
Figure 9: Western Bank displaying the location of the two study areas. Preferred (upper box) and non-preferred (lower box) juvenile Haddock habitat areas. The shaded bathymetric relief ranges from shallow (light blue) to deep (dark blue) where depth greater than 1000 m off the shelf are shaded gray.	34
Figure 10: Surficial geology of the non-preferred study area. These data were produced by marine geologists at NRCan (Fader, 2007).	36
Figure 11: Histogram displaying the bimodal depth distribution in the non-preferred study area.	36
Figure 12: Surficial geology of the preferred study area. These data were produced by marine geologists at NRCan (Fader, 2007).	37
Figure 13: Acoustic survey lines 2 to 8 for the non-preferred study area.	41
Figure 14: Diagram of the raw acoustic data processing steps using both Echoview and SAS computer software (BS = backscatter, TVG = Time-Varied Gain).	44

Figure 15: Line graph displaying an example of average backscatter strength values (dB) from the 38 kHz frequency as a function of incidence angle (degrees). Near-nadir response is defined from 3° to 6°.	51
Figure 16: Line graph displaying an example of average backscatter strength values (dB) from the 120 kHz frequency as a function of incidence angle (degrees). Near-nadir response is defined from 4° to 7°. X-axis is cut-off at 27 degrees due to extremely low off-nadir values.	52
Figure 17: Histograms of near-nadir backscatter values from the non-preferred study area.	71
Figure 18: Histograms of near-nadir backscatter values from the preferred study area.	72
Figure 19: Example of the delineation of the quantile classes (defined by the area between the blue bars) in the preferred study area. Numbers 1 to 5 indicate the quantile classes.	81
Figure 20: Results of the quantile classification of the 120 kHz near-nadir backscatter values in the preferred study area overlaid on the interpreted sediment units.	82
Figure 21: Distribution of the results of differencing the quantile classes (120 kHz – 38 kHz).	86
Figure 22: Distribution of the results of differencing the quantile classes (120 kHz – 38 kHz) summarized by sand and gravel seabeds in both study areas.	88
Figure 23: Results of the differences between quantile classifications of each frequency for the preferred study area.	89
Figure 24: Results of the differences between quantile classifications of each frequency for the non-preferred study area.	90
Figure 25: Histogram displaying the frequency of spatial continuity of points with no differences between quantile classes in the non-preferred study area.	94
Figure 26: Line graph of near-nadir backscatter values from both frequencies in the non-preferred study area along survey line 2 (surveyed from northeast to southwest). Data were smoothed using a running average of 9.	95
Figure 27: Line graph of near-nadir backscatter values from both frequencies in the preferred study area along survey line 2 (surveyed from northeast to southwest). Data were smoothed using a running average of 9.	95
Figure 28: Areas of frequency divergence within a 5-sample moving window overlaid onto the interpreted surficial geology map in the non-preferred study area.	99
Figure 29: Scatter plot of near-nadir backscatter values for the four survey lines that overlap the interpreted surficial geology map in the non-preferred study area.	102
Figure 30: Scatter plot of near-nadir backscatter values for the four survey lines that overlap the interpreted surficial geology map in the preferred study area.	103
Figure 31: Scatter plot of dual-frequency near-nadir backscatter intensities of ISUs with dominant sand sediment type in the preferred study area. The regression line is the one defined for all the backscatter data that overlapped the ISUs.	104

Figure 32: Scatter plot of dual-frequency near-nadir backscatter intensities of ISUs with dominant gravel sediment type in the preferred study area. The regression line is the one defined for all the backscatter data that overlapped the ISUs.	104
Figure 33: Scatter plot of dual-frequency near-nadir acoustic backscatter intensities symbolized by interpreted sediment units in the preferred study area.	105
Figure 34: Histogram of the backscatter residual values from the 38 kHz and 120 kHz linear regression model in the non-preferred study area.	106
Figure 35: Histogram of the backscatter residual values from the 38 kHz and 120 kHz linear regression model in the preferred study area.	107
Figure 36: Scatter plot displaying homoscedasticity of 38 kHz residuals and 120 kHz backscatter in the preferred study area.	107
Figure 37: Scatter plot displaying homoscedasticity of 38 kHz residuals and 120 kHz backscatter in the non-preferred study area.	108
Figure 38: Residual values classified into five quantile classes overlaid onto the interpreted surficial geology map.	110
Figure 39: Residual values classified into five quantile classes overlaid onto the interpreted surficial geology map.	111
Figure 40: Scatter plots displaying the relationship between morphology layers and backscatter residuals in the preferred study area. r^2 values are small for depth (0.0004), slope (0.0001) and curvature (0.000003).	113
Figure 41: Scatter plots displaying the relationship between morphology layers and backscatter residuals in the non-preferred study area. r^2 values are small for depth (0.003), slope (0.0002) and curvature (0.0027).	113
Figure 42: Histograms of discriminant function analysis input variables for the non-preferred study area.	116
Figure 43: Histograms of discriminant function analysis input variables for the preferred study area.	116
Figure 44: Matrix scatter plot of the input variables for discriminant function analysis in the non-preferred study area. Backscatter = 38_R1ADJ and 120_R1ADJ, Depth <52m = DEPTH_LS_52, Slope = SB_SLP_DEG, Curvature = MB_CURV. These plots aid in visualizing the relationship between predictors and the spread of each plot is roughly equal supporting the homogeneity of variance-covariance assumption.	117
Figure 45: Matrix scatter plot of the input variables for discriminant function analysis in the preferred study area. Backscatter = 38_R1ADJ and 120_R1ADJ, Depth = DEPTH, Slope = SB_SLP_DEG, Curvature = MB_CURV. These plots aid in visualizing the relationship	

between predictors and the spread of each plot is roughly equal supporting the homogeneity of variance-covariance assumption.	118
Figure 46: Scatter plot of the scores for the first two discriminant functions with corresponding ISUs for the preferred study area.	124
Figure 47: Scatter plot of the scores for the first two discriminant functions with corresponding ISUs for the non-preferred study area with depth <52m.	125
Figure 48: Scatter plot of the scores for the first two discriminant functions with corresponding ISUs for the non-preferred study area with depth >52m.	126
Figure 49: Map of the clustering results overlaid on the interpreted surficial geology in the non-preferred study area with depth<52m.	132
Figure 50: Map of the clustering results overlaid on the interpreted surficial geology in the non-preferred study area with depth>52m.	133
Figure 51: Map of the clustering results overlaid on the interpreted surficial geology.	133

List of Tables

Table 1: Description of the Interpreted Sediment Units (ISUs).	38
Table 2: Summary table of sand and gravel polygons on Western Bank.	39
Table 3: Summary table of the mean area (m) and percent (%) of coverage of each ISU in comparison to the whole study area.	39
Table 4: Input parameters for data processing using Echoview.	45
Table 5: Count of acoustic samples over each ISU from both acoustic frequencies. Each sample represents five acoustic pings.	70
Table 6: Skewness significance test results (S_k = standard error of skewness, z = z distribution score).	73
Table 7: Descriptive statistics of near-nadir backscatter values from both frequencies in both study areas.	74
Table 8: Descriptive statistics of near-nadir backscatter summarized by dominant sand and gravel seabed substrate types.	75
Table 9: Descriptive statistics of near-nadir acoustic backscatter of each ISU.	77
Table 10: Absolute difference in mean near-nadir backscatter values between ISUs in the preferred study area.	78
Table 11: Absolute difference in mean near-nadir backscatter values between ISUs in the non-preferred study area.	78
Table 12: Kruskal-Wallis H tests results (see section 3.4 for explanation of mean rank and Chi-square).	79
Table 13: Count of observations over each ISU from the balanced datasets in both study areas.	80
Table 14: Count of observations of quantile classes in the preferred study area.	84
Table 15: Count of observations of quantile classes in the non-preferred study area.	84
Table 16: Count of observations of differencing the quantile classes of each frequency in both study areas (Class Difference = quantile class difference value).	86
Table 17: Count of observations of differencing the quantile classes from each frequency summarized by sand and gravel seabed types (Class Difference = quantile class difference value).	87
Table 18: Count of observations of differencing the quantile classes from each frequency summarized by the ISUs in the non-preferred study area (CD = quantile class difference, F = frequency of occurrence).	91

Table 19: Count of observations of differencing the quantile classes from each frequency summarized by the ISUs in the preferred study area (CD = quantile class difference, F = frequency of occurrence).	91
Table 20: Summary of the spatial continuity of points from quantile class differencing in the preferred study area. Summarized by no difference, higher quantile class for the 38 kHz frequency (+ 38), and higher quantile class for the 120 kHz frequency (+ 120).	93
Table 21: Summary of the spatial continuity of points from quantile class differencing in the non-preferred study area. Summarized by no difference, higher quantile class for the 38 kHz frequency (+ 38), and higher quantile class for the 120 kHz frequency (+ 120).	93
Table 22: Descriptive statistics of the along-track distances (meters) of high and low near-nadir backscatter areas in both study areas. High and low backscatter values were defined by the following cut-off values: -11 dB (120 kHz) and -14 dB (38 kHz) in the non-preferred study area and -10 dB (120 kHz) and -14 dB (38 kHz) in the preferred study area (BS = backscatter).	97
Table 23: Percentages of samples with frequency divergences based on certain window sizes.	98
Table 24: Non-preferred study area Wilcoxon matched-pairs signed-ranks test of near-nadir backscatter values from each frequency for each ISU (see section 3.6 for explanation of sum of ranks used, mean rank, sum of ranks, and Z).	100
Table 25: Preferred study area Wilcoxon matched-pairs signed-ranks test of near-nadir backscatter values from each frequency for each ISU (see section 3.6 for explanation of sum of ranks used, mean rank, sum of ranks, and Z).	100
Table 26: Descriptive statistics of residual values.	109
Table 27: Count of observations of residual quantile classes for the preferred study area.	112
Table 28: Count of observations of residual quantile classes for the non-preferred study area.	112
Table 29: Skewness significance test for discriminant function analysis input variables.	115
Table 30: Correlation matrix of input variables for discriminant function analysis.	118
Table 31: Test of equality of group means of the discriminant function analysis for the preferred study area.	119
Table 32: Standardized discriminant function coefficients.	120
Table 33: Correlation between predictor variables and discriminant functions.	121
Table 34: Wilks' Lambda results for the preferred study area.	122
Table 35: Eigenvalues and explained variance of the discriminant functions for the preferred study area.	122
Table 36: Mean discriminant scores for each function.	123
Table 37: Classification function coefficients for the preferred study area.	127
Table 38: Classification function coefficients for the non-preferred study area.	127

Table 39: Classification results of the discriminant function analysis for the preferred study area.	128
Table 40: Classification results of the discriminant function analysis for the non-preferred study area with depth <52m.	129
Table 41: Classification results of the discriminant function analysis for the non-preferred study area with depth >52m.	129
Table 42: Classification accuracies for the supervised classification results.	130
Table 43: Distances between final cluster centers for the k-means classification in the preferred study area.	131
Table 44: Distances between final cluster centers for the k-means classification in the preferred study area with depth <52m.	132
Table 45: Distances between final cluster centers for the k-means classification in the preferred study area with depth >52m.	132
Table 46: k-means class modes for each ISU in both study areas.	134
Table 47: Summary of mean near-nadir backscatter values within each study area summarized by frequency and dominant substrate type.	135
Table 48: Summary of general results of each ISU. QC = Quantile classification results, QCD = quantile classification differencing results, DFA (%) = supervised classification accuracy percent. QC and QCD values (Y = yes and N = no) indicate whether the ISU followed the typical trend of higher near-nadir backscatter from gravel substrate then sand substrate. The Wilcoxon values indicate whether the mean near-nadir backscatter values from each frequency was significantly different for each ISU.	135

List of Abbreviations and Symbols

BS – Backscatter

dB – Decibel

CTD – Conductivity-Temperature-Depth

DFO – Department of Fisheries and Oceans

GIS – Geographic Information System

GPS – Global Positioning System

GSC – Geological Survey of Canada

ISUs – Interpreted Sediment Units

GRS – gravel ripples – short wavelength

GRL – gravel ripples – long wavelength

GRI – gravel ripples incised

GL – gravel lag

GH – gravel hummocky

Sand – sand

SM – sand with megaripples

SG – sand to sandy gravel

SB – sand with boulders

kHz – Kihertz

MBES – Multibeam Echo Sounder

MHz – Megahertz

pps – Pings per Second

SA – Area Backscattering Coefficient

SBES – Single Beam Echo Sounder

SE – Standard Error

SNR – Signal to Noise Ratio

SSS – Sidescan Sonar

Sv – Volume Backscatter Coefficient

TVG – Time-Varied Gain

List of Appendices

Appendix A: Interpreted Sediments Units (ISUs)	165
Appendix B: Mann-Whitney U Test Results	166
Appendix C: Quantile Classification Maps	170
Appendix D: Frequency Differences Based on Classified Backscatter Values	173
Appendix E: Frequency Differences Based on Continuous Backscatter Values	177
Appendix F: Multivariate Analysis	185
Appendix G: Balanced Backscatter Data	187

1. Introduction

1.1 Context

The world's oceans play a crucial role in human life as they provide a significant amount of renewable (e.g. waves and tides) and non-renewable (e.g. oil and gas) energy resources, essential food resources provided by fisheries, crucial transportation and shipping routes, a myriad of recreation opportunities, as well as playing a major role in the earth's climate. The seabed is a crucial part of the marine environment as it provides vital habitat for demersal (bottom dwelling) fish and benthos (organisms living on or in the seabed). Nation states are placing considerable effort in developing detailed maps of the continental shelf seabeds over the next few decades. Much of the motivation for nation states to map the seabed arose from the United Nations Convention on the Law of the Seas (UNCLOS) in 1982. With the signing of this convention, nation states are able to legally extend the limits of their jurisdictional authority over territorial waters and the continental shelf, including the management of marine natural resources (UNCLOS, 1982). In this context, seabed mapping plays an important role as it is used to establish the extent of a nation's continental shelf and determine governing authority. Along with the establishment of authority over territorial waters and the continental shelf comes a responsibility to scientifically manage the resources in the area.

The management of marine resources requires an ecosystem-based approach whereby natural regions are identified and mapped over a range of spatial scales (Durand et al., 2006, Anderson et al., 2008). In this context, we must be careful to develop objective methods and techniques for gathering and analyzing data on the marine environment, as there is an ever-growing demand to map the seabed at varying

scales. Such efforts are necessary for a number of reasons: (1) in order to manage and use resources in a sustainable fashion without exerting too much stress on the marine environment; (2) to protect sensitive species, habitats, and ecosystems; (3) to monitor environmental changes; and (4) to assess the impacts of anthropogenic disturbances. Hence, objective and repeatable scientific methods for collecting data, deriving relevant information, and inferring knowledge should provide a framework for better management and informed decision-making concerning the marine environment. These steps can be illustrated in a hierarchical system, or knowledge pyramid, representing the structural relationship of data, information and knowledge (Figure 1).



Figure 1: Knowledge pyramid representing levels of understanding from data to knowledge. Adapted from Poore and Chrisman (2006).

Mapping the seabed at multiple scales in support of ecosystem-based scientific management of marine resources requires a new generation of assessment tools and techniques (Anderson et al., 2008). Collecting and mapping data of physical seabed characteristics is almost entirely performed from the sea surface using underwater acoustic-based remote sensing systems. Acoustic waves do not attenuate much in water, in comparison to electromagnetic energy, and typically travel at an average

propagation speed of 1500 m/s, four to five times faster than in the air. As a result, underwater acoustics techniques are the most practical way to transmit data underwater (Lurton, 2002, Buchanan et al., 2004). Acoustic systems can ultimately carry information through thousands of meters of water and the effective range of the acoustic system primarily depends on the frequency (Lurton, 2002). Other techniques such as sediment grabs and underwater photographs and videos can provide direct and detailed observations of seabed substrate (i.e. the sediment that constitutes the seabed) and topography. However, these techniques require that the device be directly on or near the seabed and only provide coverage of a small area. Acoustic remote sensing, on the other hand, provides much more coverage of the seabed in a systematic and repeatable manner. Seabed mapping procedures have developed at a rapid pace due to the advancement of underwater acoustic technology systems and survey techniques (Kenny et al., 2003). These systems have led to the design of new approaches for mapping marine benthic habitats, for instance, combining high-resolution acoustic data with direct observations of the seabed (i.e. ground truthing) using sediment samples, photographs and videos (Kostylev et al., 2001, Kloser et al., 2002, Roberts et al., 2005, Brown and Collier, 2008).

Underwater echo sounders are effective tools for collecting data from the seabed based on acoustic energy directed downwards toward the seabed. Echo sounders have sensors (i.e. transducers) that produce mechanical vibrations, which project sound waves (i.e. pressure waves) into the water. These sound waves are emitted at a given frequency, propagate through the water column, and ensonify a circular area on the seabed or other objects in its path (e.g. fish, vegetation). Part of the acoustic signal is absorbed into objects, part of it is scattered away from the sensor, and part is reflected or scattered back towards the acoustic system (Lurton, 2002). This returned acoustic

energy from different objects is recorded by the system and is referred to as acoustic backscatter, or the returned echo or acoustic signal. These systems then convert the acoustic energy return into digital form (analog to digital conversion) for computer use, which allows scientists to read, interpret, and understand the data. Backscatter data are used to map and classify seabed substrate characteristics based on models that consider how substrate influence the shape and intensity of returned acoustic signals (Jackson and Briggs, 1992; Sternlicht and de Moustier, 2003). The shape (i.e. width) and intensity (i.e. peak) of the returned acoustic signal or response curve can be used to derive information about the roughness (a measure of the texture and irregularities of the seabed) and hardness (a measure of the acoustic impedance contrast at the water-seabed interface) of the seabed (Courtney et al., 2005).

Mapping the physical and biological characteristics of the seabed based on the collection, conditioning (i.e. processing), and classification of acoustic seabed backscatter data is called Acoustic Seabed Classification (ASC). More specifically, ASC objectively classifies seabed types based on remotely sensed acoustic measurements of the seabed's physical and biological attributes (Simard and Stepnowski, 2007; Anderson et al., 2008). In its simplest form, ASC involves collecting acoustic data from the seabed as a continuous variable, and then turning these data into information classes that reflect physical seabed characteristics. Seabed types are typically defined based on both the grain size of the sediment as well as the roughness and morphology. An important aspect of ASC is the different classification techniques used to derive the specific sets of clusters or classes that represent seabed types. There are many classification methods used in ASC (Simard and Stepnowski, 2007) and these can be broken down into supervised and unsupervised methods. Supervised classification is a technique that 'trains' the algorithm to classify a dataset based on *a priori* information about the typical

acoustic signatures or statistical characteristics that certain physical seabed types produce. Unsupervised classification is a technique that classifies the data into groups without any prior knowledge of the envisioned seabed. Instead, this technique is based on statistical clusters of the data. Regardless of which classification technique is used, ASC is able to derive information about the seabed by grouping or clustering similar measures of one or more attributes. For example, depth, grain size, and roughness are all physical attributes that acoustic backscatter can measure and the appropriate classification can theoretically sort their values into relatively homogenous groups despite the complexity of seabed scattering.

Most approaches used for acoustic seabed classification and mapping use a single acoustic frequency approach (e.g. Anderson et al., 2002, Courtney et al., 2005). Anderson et al. (2002) used a commercial single frequency acoustic seabed classification system (QTC View) to classify marine benthic habitats. This classification system was trained, based on geo-referenced calibration sites that represent typical backscatter responses for different seabeds, to classify a set of seabed types over a broader area using real-time classification. This paper effectively illustrates the usefulness of ASC techniques to delineate seabed habitats. Courtney et al. (2005) used both a single beam echo sounder and a sidescan sonar (see Chapter 2 for a review of these systems) operating at the same frequency to collect acoustic data. The authors derived two metrics from the acoustic backscatter signal, namely shape and intensity, from the single beam system and the mean and variance from the sidescan, in an unsupervised classification. Results revealed the challenges of relating statistical clusters to previously defined surficial geology classes. Although ASC had been typically approached using single frequency backscatter data, there have been a few studies (e.g.

Kloser et al., 2002) that used multi-frequency backscatter data to improve ASC attempts (details of existing multi-frequency ASC research are outlined in section 1.2).

Theory relating backscatter to certain seabed types is based on models, which quantify the relationship between physical seabed properties and the corresponding backscatter signal's shape and intensity (Jackson and Briggs, 1992, APL94, 1994, Sternlicht and de Moustier, 2003). In addition to relating seabed types to specific acoustic signatures of backscatter produced by the physical properties of the seabed, these models also consider the specific frequency used by the acoustic system. Hence, the amount of absorption, attenuation, and reflectivity of the acoustic signal for various seabed types is not only governed by physical seabed properties, but also by the particular frequency used by the system (McKinney and Anderson, 1964, Galloway and Collins, 1998, Kloser et al., 2001, Lurton, 2002). In principle, lower acoustic frequencies travel further and penetrate deeper into the seabed whereas higher frequencies resolve smaller spatial features such as sediments and surficial roughness (Galloway and Collins, 1998, Lurton, 2002). Hence, the acoustic frequency plays a crucial role on the shape and intensity of backscatter. Seabed mapping applications typically use frequencies greater than 10 kHz as the acoustic response from frequencies in this range do not penetrate far into the seabed (tens of centimetres depending on the seabed type) and therefore contain information on the topography and substrate of the seabed (Anderson et al., 2006), however range is limited to less than 10 kilometres (Lurton, 2002). In contrast, sediment profilers and seismics use much lower frequencies to penetrate deeper into the seabed, using frequencies less than 10 kHz (much less for seismics), and have greater ranges travelling up to hundreds or even thousands of kilometres depending on the frequency (Lurton, 2002).

ASC is ultimately a very effective technique for exploring and understanding the seabed at a variety of scales. It is important to consider that acoustic seabed backscatter varies as a function of the specific frequency used. This potential added information for detecting and classifying seabed types could improve efforts to identify and define the distribution of marine benthic habitats of benthos and demersal fish species, as these habitats are believed to be primarily determined by substrate type (Gregory and Anderson, 1997, Kostylev et al., 2001).

1.2 Research Problem

While the use of multiple acoustic frequencies is seen as one of the main future research paths for acoustic seabed classification (Anderson et al., 2008), the analysis of backscatter from multiple acoustic frequencies for the purposes of acquiring additional information to improve ASC remains largely unexplored. Few studies have explored the use of multiple acoustic frequencies for classifying surficial geology and differentiating seabed habitats (Galloway and Collins, 1998, Kloser et al., 2002, Fossa et al., 2005, Riegl and Purkis, 2006, Chakraborty et al., 2007).

Galloway and Collins (1998) used a commercial ASC system (QTC View) to classify surficial geology using 38 kHz and 200 kHz acoustic frequencies. This study concluded that dual frequency data, when used in conjunction with single frequency classification results, provided additional insight into surficial geology. They also concluded that the ability of the lower frequency to penetrate deeper into the seabed provided information about the physical properties of the substrate volume. In contrast, the higher frequency detected the immediate water-seabed interface and therefore was more influenced by surface scattering and could detect smaller particles related to sand and mud substrate (see sections 2.2, 2.3, and 2.4 for explanation of both surface and

volume scattering). Kloser et al. (2002) used both user developed and proprietary software to process acoustic data and classify seabed biotopes of single beam (12 kHz, 38 kHz, 120 kHz, and combined 12 and 38 kHz) and multibeam acoustic datasets. Classification results were compared to ground-truthing and the combined frequency dataset proved to have lower cross-validation error. Fossa et al. (2005) provided an assessment of how backscatter responses from corals varied using a number of acoustic frequencies (18 kHz, 38 kHz, 120 kHz, and 200 kHz). Riegl and Purkis (2005) used a commercial acoustic seabed classification system (QTC View) and found that a lower frequency (50 kHz) carried more information about the hardness of the substrate and therefore resulted in a two-class outcome, namely soft and hard seabed. The higher frequency (200 kHz) primarily carried information about the roughness of the seabed and was able to distinguish rough and flat seabeds. Overlaying classification maps from both frequencies resulted in a four-class map containing soft, hard, rough, and flat seabeds. Finally, Chakraborty et al. (2007) highlighted how backscatter from a higher (210 kHz) frequency was a function of the immediate water-seabed interface while a lower frequency (33 kHz) penetrated deeper into the seabed and therefore encounters different subsurface sediment layer compositions resulting in volume scattering. They also concluded that the 210 kHz frequency was more effective for discriminating between sediment types as the differences between backscatter values of fine and coarse grain sediment was greater. These results reflect those of Galloway and Collins (1998) and Riegl and Purkis (2005). These papers have empirically shown how added information can be obtained from using multiple frequencies and how this information can improve ASC attempts.

However, there are several issues with the existing research using multiple frequencies. First, the research highlights the differences in acoustic backscatter from

different frequencies, but typically does not attempt to combine the multi-frequency backscatter into a multivariate acoustic seabed classification scheme. Second, if multi-frequency backscatter is combined for seabed classification, it often does so using commercial black-box ASC software such as RoxAnn, EchoPlus, QTC VIEW, and QTC IMPACT, which allow for a limited understanding of data properties in terms of processing and interpretation that may not be replicable (Anderson et al., 2008). Therefore, the need for a new method that analyzes the effectiveness of combining multi-frequency backscatter, outside of commercial software, is required in order to provide more transparency of the data processing and statistical methods used in ASC. This will allow researchers to understand seabed scattering processes at an elementary level that permits proper interpretation of results (Anderson et al., 2008). Finally, it would be useful to identify areas of frequency divergences and the spatial scales at which this occurs, which has not been the specific focus of previous studies.

1.3 Research Hypothesis and Research Questions

The hypothesis followed in this project is that combining near-nadir acoustic backscatter from two frequencies will provide additional information on seabed substrate properties, consequently improving our understanding of seabed types, and therefore improve upon ASC attempts. Near-nadir backscatter is measured from the returned acoustic signal based on a certain period from when the leading edge of the signal first encounters the seabed until the signal becomes fully embedded at near normal incidence. In other words, near-nadir is the initial spread of the acoustic beam closest to orthogonal with the seabed (see chapter 2 for further explanation). Backscatter is dependent upon a number of factors including changes in depth and ship motion with corresponding variations on beam pattern and incidence angles, sound speed variations

in the water column, as well as the configuration of the acoustic system, including its frequency, pulse length, beam angle, and beam pattern. The backscatter data used here went through a data processing procedure that compensated for changes in depth, sound speed variations, and the system configurations (pulse length and beam angle) remained relatively constant (see section 3.3 for further explanation). In addition, ship motion and sea-state conditions were minimized by only surveying in calm conditions, however, ship motion was not compensated for and will be a source of measurement error (see section 3.3). Therefore, the only changing constants are the physical seabed properties, frequency, and the incidence angle at which the backscatter signal is analyzed. Standardizing the backscatter data to represent the near-nadir component (see section 2.3) leaves only physical seabed types and frequency as the two determining factors on the backscatter. The backscatter data used in this research were acquired using single beam normal incidence echo sounders operating at 38 kHz and 120 kHz frequencies collected on Western Bank, part of the Scotian Shelf (see chapter 3 for further descriptions of the study area and data collection methods). Seabed backscatter varies as a function of seabed substrate properties and the acoustic frequency used, therefore additional information for delineating seabed substrate types should result from using multi-frequency backscatter. Different responses are expected from different acoustic frequencies based on the amplitude and wavelength of each frequency. These factors determine the ability of a sound wave to penetrate the seabed, primarily as a function of particle size and porosity, and to detect roughness at scales comparable to the wavelength (see chapter 2 for further discussions of underwater acoustic backscatter, acoustic frequency, and seabed substrate properties). Consequently, and as suggested by previous studies, a lower frequency should detect features that a higher frequency will not and vice versa. It is also anticipated that the use

of multiple frequencies should be most effective when there are transition zones between seabed substrate types or the seabed substrate distribution is heterogeneous, as a mixture of sediments should best be detected with more than one frequency.

Six research questions, grouped into four categories, are asked. The first category (questions 1 and 2) is concerned with the statistical summary and analysis of both single and dual frequency near-nadir backscatter as a function of frequency and seabed substrate types. The second category (question 3) examines the spatial patterns of frequency differences. The third category (questions 4 and 5) focuses on relating classified near-nadir acoustic backscatter to the seabed substrate types. The fourth and final category (question 6) addresses the question of whether or not dual-frequency classification improves upon single frequency classification. The research questions for this project are:

1. What are the univariate (single frequency) statistical characteristics of near-nadir acoustic backscatter data from each frequency independently and do they differ?
2. How do univariate classifications of near-nadir acoustic backscatter values of each frequency relate to seabed substrate types?
3. Where do the major frequency differences occur and what are the spatial scales at which these frequency differences take place?
4. What are the multivariate (dual-frequency) statistical characteristics of near-nadir acoustic backscatter data from both frequencies?
5. How do multivariate classifications of dual-frequency near-nadir acoustic backscatter values relate to seabed substrate types?
6. Does dual-frequency near-nadir acoustic backscatter improve our ability to classify seabed substrate types compared to a single acoustic frequency?

1.4 Project Objectives

The goal of this research is to assess differences and similarities in dual-frequency near-nadir acoustic backscatter from the seabed, and to determine how the combination of dual-frequency backscatter compares to single-frequency backscatter classification of seabed substrate characteristics. The spatial scales at which frequency differences occur are identified. Improvements resulting from dual-frequency classification in comparison to single-frequency classification attempts are assessed. Objective and repeatable methods will be used for assessing the characteristics of single and dual-frequency near-nadir backscatter, and for the classification of backscatter values in comparison to seabed substrate types. Specific objectives are:

1. to analyze the statistical characteristics and relationships of near-nadir acoustic backscatter values as a function of frequency and seabed substrate characteristics;
2. to identify where frequency differences occur on the seabed and at what spatial scale;
3. to classify near-nadir backscatter values in an objective manner for each frequency individually, and in combination, and assess classification accuracy by comparing results to previously defined seabed substrate types;
4. to determine if dual-frequency classifications improve upon single frequency classification.

1.5 Methods

Calibrated acoustic backscatter data from both 38 kHz and 120 kHz acoustic frequencies were collected and processed as explained in sections 3.2 and 3.3. The author was not part of the data collection process, rather the acoustic data were collected by scientists at Department of Fisheries and Oceans Canada (DFO). The data were collected along transect lines using a single-beam echo sounder (SBES). The

backscatter data were processed to represent standardized near-nadir backscatter responses, where frequency dependence of backscatter is deemed highest (see sections 2.2, 2.3, 2.4, and 3.3 for further explanation). Existing interpreted surficial geology maps of the study areas were acquired from Natural Resources Canada (NRCan), which contained interpreted sediment units (ISUs) representing physical seabed classes based on grain sizes and geomorphology characteristics (see section 3.1 for further explanation). Specific, repeatable, and objective statistical techniques and analyses were performed (see Chapter 3) using these data in order to address the specific research objectives. Results from this study were also presented at several national and international scientific conferences (SKI 2009 Spatial Knowledge and Information Conference, GeoHab 2009 Marine geological and Biological Habitat Mapping, CoastGIS 2009 Symposium and on GIS and Computer Cartography for Coastal Zone Management, GAC NL 2010 Geological Association of Canada NL Branch Annual Technical Meeting) before being published.

Generally, the analytical methods can be broken down into univariate analysis, using backscatter from one frequency, and multivariate analysis, using backscatter from two frequencies. In the univariate case, descriptive and inferential statistical techniques were used to understand the typical trends of acoustic backscatter data as a function of frequency and seabed substrate characteristics. Univariate classification techniques were used to statistically group the data and relate the results to seabed substrate types.

In the multivariate case, statistical techniques, such as regressions and correlations, were used to help identify relationships between backscatter data of the two frequencies. Statistical analyses were used to relate dual-frequency backscatter characteristics to specific seabed types. Areas of frequency divergences between dual-frequency backscatter values were identified along with the spatial scales at which this

occurred. Multivariate classification methods used single and dual-frequency backscatter and additional morphology variables (depth, slope, curvature) with emphasis placed on relating results to seabed types on the Scotian Shelf.

1.6 Thesis Layout

Chapter 1 presented research context, the research problem this thesis addresses, hypothesis and questions, objectives, and an outline of the method. Chapter 2 summarizes appropriate background information and literature relevant for this thesis, such as underwater acoustic systems, acoustic backscatter, acoustic frequencies, and acoustic seabed classification techniques. Chapter 3 outlines research methods, including a detailed description of the study area, data collection, and processing methods. Chapter 3 also describes the different univariate and multivariate statistical techniques and analyses used to address the objectives and answer the research questions. Chapter 4 presents results of the analyses of backscatter data. Chapter 5 provides a discussion of results, relating specific findings to previous multi-frequency studies and providing possible theoretical explanations for specific trends and patterns in the data. Chapter 6 provides conclusions that were drawn from this research and outlines possible future research directions, and discusses the implications of this research for decision makers and real-world applications.

2. Literature Review

This chapter summarizes some of the relevant existing knowledge found in the literature that pertains to this study. First, section 2.1 outlines the field of Acoustic Seabed Classification (ASC) and its typical practices and methods. ASC involves the collection, processing, and classification of surficial seabed classes based on acoustic measurements of physical and biological properties. Understanding acoustic seabed backscatter is crucial in the field of ASC, and therefore the important concepts and knowledge from the acoustic seabed backscatter literature are highlighted in section 2.2. The characteristics of acoustic signals collected using single beam echo sounders are discussed in section 2.3. Finally, an important part of this project is the influence of acoustic frequency on seabed backscatter, and therefore acoustic frequencies are reviewed in section 2.4.

2.1 Acoustic Seabed Classification (ASC)

ASC follows a set of systematic steps that include the collection, conditioning, and classification of acoustic backscatter data (Anderson et al., 2008) and these steps are shaped by the classification objective (Simard and Stepnowski, 2007). It is important to consider how the survey design will account for spatial and temporal variability, and this should be determined by the specific seabed classification scheme (Anderson et al., 2008). Specifically, both transect line spacing and seabed directionality are important considerations in survey design. Transect spacing will influence the extent of seabed that an acoustic system can cover and the sailing direction of the ship will determine how seabed morphology (i.e. slopes and ridges) is represented in the data.

Data conditioning is a crucial part of ASC and it refers to the processing stage of the backscatter data that occurs after data collection and before classification. Removing corrupt data, increasing the signal-to-noise ratio (SNR), calibration, representing data in physical units, and consideration of platform movements are all important steps in the data conditioning stage (see Simard and Stepnowski (2007) for a detailed discussion of these steps). The quality of the data must be assured by removing bad pings that are a result of electrical or acoustic interferences. Similarly, data conditioning involves removing pings with low signal-to-noise ratio (SNR) or averaging pings to improve SNR. Although averaging or binning pings reduce spatial resolution, it is justified by the high degree of overlap between footprints. Calibration of echo sounders is also necessary for consistent backscatter measurements (see Simmonds and MacLennan (2005) for a detailed discussion of calibration methods). Calibration of an acoustic system involves the consideration of source level (acoustic level or power delivered by a transducer), pulse length (length of time that an acoustic wave is transmitted), beam pattern (refers to the change in sensitivity in reference to direction), receiver sensitivity (quantified quality of electro-acoustic conversion), and time-varied gain (TVG) (compensation of spreading and absorption loss with range) (Simard and Stepnowski, 2007). Representing backscatter data in absolute physical units, such as area backscattering strength in decibels (dB), is also an important step in the data conditioning process. Finally, corrections for platform movement (pitch, heave, roll, and yaw) must be considered as they affect the incidence angle and range of echoes.

Another important aspect of the data conditioning stage is how backscatter characteristics (i.e. shape and intensity) are used to determine the nature of the seabed (i.e. seabed roughness and hardness) based on sound-scattering theory (see sections 2.2, 2.3, and 2.4 for more information on sound-scattering theory, roughness and

hardness metrics, and how seabed substrate characteristics influence backscatter). However, there is no simple relationship between backscatter and physical seabed properties as acoustic echoes from the seabed are complex and have a low signal-to-noise ratio (Anderson et al., 2008).

The final step in ASC is the objective classification of acoustic backscatter data into discrete classes based on similar acoustic responses (Collins and Rhynas, 1998). These groups are interpreted to represent specific seabed classes. Attempts to categorize natural processes or phenomena in an environment are aided by objective methods used to classify or group discrete subsets of individual occurrences, based on similar measures of their physical and biological attributes (Legendre and Legendre, 1998, Shaw, 2005). Such techniques allow researchers to explore and simplify their datasets, extract thematic information, and make inferences about specific characteristics of their study area. This type of analysis is often referred to as classification (Shaw, 2005). The two types of classification, supervised and unsupervised, outlined in section 1.1 plays a crucial role on the classification results. Supervised techniques are dependent on the accuracy of the defined classes in the training dataset that are based on a knowledgeable description and classification of the seabed, whereas unsupervised approaches are dependent on statistical variability, which is determined by seabed diversity (Simard and Stepnowski, 2007). Classification results can be mapped and compared to ground-truthing in order to get an indication of classification accuracy and validation.

ASC is becoming common practice for seabed classification, and in comparison to traditional seabed geology mapping techniques, is much more cost-effective (Galloway and Collins, 1998, Anderson et al., 2008). Frequencies greater than 10 kHz are used in ASC as they typically do not penetrate the seabed more than one meter and

best represent the topography and substrate of the water-seabed interface of interest. Underwater acoustics and data-processing technologies are powerful tools, and when used in the context of ASC, allows for the classification and mapping of physical and biological attributes of the seabed across a continuum of spatial scales. This expands our knowledge of the marine environment and supports ecosystem-based management of marine habitats and resources when linked with habitat information, population productivity, and biodiversity (Anderson et al., 2006; Cogan et al., 2009).

2.2 Acoustic Seabed Backscatter

It is important to understand how the physical and biological properties of the seabed are embedded in acoustic backscatter through the shape and intensity of returned acoustic signals. A typical acoustic signal from the seabed rises to peak amplitude (i.e. intensity) by the time the entire pulse has encountered the seabed, which is a function of the pulse length. The signal declines over time (i.e. increasing incidence angles) to a tail that is irregular in shape and eventually declines to zero (see the graph in Figure 2 as an example of a typical return signal from the seabed displaying both the intensity and shape as a function of time). The shape of backscatter refers to distortions in the shape and duration of the original acoustic pulse (Stemlich and de Moustier, 2003). The intensity of backscatter simply refers to the amplitude or strength of the returned acoustic echo. The physical and biological nature of the seabed causes variations in the shape and intensity of returned acoustic echoes.

In the context of this work, returned acoustic signals represent only the sound waves that have been reflected or scattered back to an echo sounder from the seabed. Specular reflection occurs when the seabed reflects sound in the specular direction at an angle equal to the angle of the incident sound as a result of impedance at the water-

seabed interface (Jackson and Richardson, 2007). Specular reflection is more dominant using normal or near-nadir incidence angles as opposed to oblique incidence angles. Scattering occurs primarily as a result of surface roughness at scales comparable to the acoustic wavelength and scatters sound waves in all directions including back to the source. Hence, backscatter (including both reflection and scattering processes) results from surface roughness, impedance contrast, and volume inhomogeneities that cause scattering from within the substrate volume (Lurton, 2002, Jackson and Richardson, 2007).

The most important physical characteristics of the seabed that cause variations in the physical properties of backscatter signals are grain size, porosity, and roughness (McKinney and Anderson, 1964, Clay and Leong, 1974, Parrott et al., 1980, Jackson and Briggs, 1992, Sternlicht and de Moustier, 2003). Sediment types are classified based on grain size and are often referenced to a scale (e.g. the Wentworth scale). For example, according to the Wentworth scale, clay ($0.4 \mu\text{m} - 3.9 \mu\text{m}$) has a smaller grain size than sand ($62.5 \mu\text{m} - 2 \text{ mm}$), and sand has a smaller grain size than gravel ($2 \text{ mm} - 64 \text{ mm}$), and so on. Of course, grain size also determines porosity. Correlating acoustic properties with grain size has been previously conducted in a number of studies (e.g. McKinney and Anderson, 1964, Nafe and Drake, 1964). The seabed also has varying degrees of roughness. Roughness at scales larger than the acoustic footprint add little to the predictive power of acoustic backscatter for determining seabed types, and therefore only roughness at the same order magnitude as the acoustic footprint (centimetre to tens of meters) is of concern (Jackson and Richardson, 2007). Figure 2 illustrates how surface roughness increases the duration and distorts the shape of a returned acoustic signal.

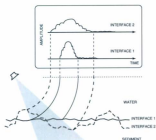


Figure 2: Typical acoustic signal returns from a smooth surface (interface 1) and a rough surface (interface 2). These graphs illustrate the influence of roughness on the shape of the returned acoustic signal (source: Sternlicht and de Moustier, 2003).

Heterogeneities in the physical structure of the seabed (i.e. surface roughness and grain size variations) cause discontinuities in the physical properties of sound waves by intercepting and reflecting or scattering a portion of the incident acoustic energy (Jackson and Richardson, 2007). More specifically, this process creates a contrast in the impedance of sound at the water-seabed interface (Jackson and Richardson, 2007). The impedance is a measure of the seabed's ability to resist the penetration and travel of sound waves through it, in other words seabed hardness. A higher impedance contrast means more sound energy is reflected or scattered from the water-seabed and less energy penetrates into the substrate volume. The density and porosity of substrate primarily influences the impedance contrast of the seabed, and these properties are related to grain size (Mank et al., 2006, Jackson and Richardson, 2007). Density is defined as the mass per unit volume and porosity is the ratio of the total volume of water

in sediment to the total sediment volume. Empirical studies (Hamilton and Bachman, 1982; Manik et al., 2006) had found that density was inversely related to porosity. A smooth and soft surface such as mud has a low impedance contrast, and therefore may not produce a strong backscatter signal (Jackson and Richardson, 2007). Therefore, substrate with larger grain sizes and greater density, generally will have higher acoustic backscatter values, and vice versa (Manik et al., 2006).

There are two types of seabed backscatter: surface (i.e. backscatter from the immediate water-seabed surface) and volume (i.e. backscatter from within the seabed). The total acoustic backscatter strength is the sum of both (Lurton, 2002; Sternlicht and de Moustier, 2003). Hence, a specific sampled area will produce a unique acoustic signature based on the contribution of both surface and volume scattering (Lurton, 2002), as is illustrated in Figure 3. Acoustic impedance contrast determines the amount of backscatter from the immediate water-seabed interface and the amount of sound energy penetrated into the substrate volume. The physical composition (i.e. heterogeneities) of the substrate volume influences volume scattering (Lurton, 2002). Seabed roughness has more influence on surface scattering from the water-seabed interface at scales comparable to acoustic wavelengths (Lurton, 2002; Sternlicht and de Moustier, 2003). Volume scattering is dominant in softer seabeds (i.e. smaller grain sizes) and surface scattering and roughness are dominant in harder seabeds (i.e. larger grain sizes) (Jackson and Briggs, 1992; Sternlicht and de Moustier, 2003). Sound also interacts differently with the seabed at varying angles of incidence (i.e. *time*), and backscatter intensity is generally higher near-nadir (i.e. near-perpendicular) as opposed to angles off-nadir (off-perpendicular).

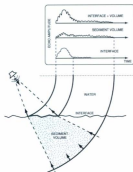


Figure 3: Acoustic backscatter response showing both the surface (i.e. interface) and volume scattering of a returned acoustic signal as a function of time (source: Sternlicht and de Moustier, 2003).

To summarize, the sum of both volume and surface reflection or scattering influence backscatter and is a result of the geoaoustic properties of the seabed (i.e. sound speed, impedance, and attenuation). These geoaoustic properties are determined by the physical properties of the seabed (i.e. roughness, grain size distribution, porosity, and density), which create heterogeneities that cause sound to be reflected or scattered (Jackson and Richardson, 2007). The shape of the acoustic signal is primarily influenced by roughness of the seabed and volume scattering. The intensity of the acoustic signal is influenced by the hardness of the seabed. The acoustic hardness of the seabed is determined by the acoustic impedance contrast, which is a function of grain size.

2.3 Single Beam Echo Sounders

Three primary types of acoustic technology systems are commonly utilized for conducting surveys on the seabed (see Kenny et al. (2003) for a review of these systems), namely multibeam echo sounders (MBES), sidescan sonar (SSS) and single beam echo sounders (SBES). Sidescan sonar is more of an acoustic imaging device that provides wide-area pictures of the seabed reflecting the texture and roughness at incidence angles ranging from 40° to 85° (Kloser, 2007). However, SSS usually do not provide georeferenced bathymetric information (NOAA, 2009) and interpretation of seabed roughness and texture are based on visual subjective interpretation more than objective statistical and image processing techniques (Anderson et al., 2008). For these reasons, SSS is not the preferred type of system for this specific project or for ASC.

There are advantages and disadvantages to using a SBES for seabed classification. A primary disadvantage is that a SBES does not provide full coverage of the seabed like a MBES or SSS does, and therefore inter-track interpolation is required if complete coverage is needed. On the other hand, the use of a SBES has many advantages: on-axis calibration, use of multiple acoustic frequencies during a single survey, relatively low acquisition costs, efficient data processing due to low data volumes, relative ease of understanding, and standard processing techniques (Anderson et al., 2008). More importantly, "... SBESs have the advantage of ensonifying the seabed from the same incident beam, and providing the same acoustic information from the unresolved footprints" (Anderson et al., 2008, p.1005).

When comparing backscatter values from different study areas, times of acquisition, or systems, there must be a certain amount of quality control in the process. In other words, backscatter values from different systems or frequencies must be calibrated in order to directly compare them. Calibration is "... conducted to determine

the correct value of the scale reading of an instrument, by measurement or comparison with a standard" (Simmonds and MacLennan, 2005, p.110). There are internationally accepted standards for calibrating SBES systems (Foote et al., 1987).

The design and configuration of the transducer and acoustic system influences the shape and strength of the desired acoustic signal. The transducer shapes the acoustic signal into an approximate cone directed vertically downwards toward the seabed (Figure 4), although the transducer also generates side-lobes that make the signal more complex in practice (Foster-Smith and Sothoran, 2003). The acoustic beam has certain dimensions depending on the specific configuration of the acoustic system. The beamwidth and pulse length shapes the dimensions of the beam. Beamwidth and range, as well as operational considerations that include vessel speed and sampling rate (i.e. pings per second), will determine the size of the acoustic footprint (i.e. size of the area on the seabed ensounded by the acoustic beam). The acoustic system also controls the intensity and frequency of the emitted pulse of sound.

As an acoustic beam from a SBES reaches the seabed it constantly spreads across and penetrates the seabed until the signal eventually dissipates. The acoustic beam ensounds the seabed in a more or less circular fashion. The ensounded area of a SBES at normal incidence on a flat seabed can be illustrated as an expanding disk between contact time t_c and t_p (Kloser, 2007). At this period, the acoustic beam has been fully embedded in the seabed and the signal has reached its peak backscatter signal, this is a function of the pulse length and is considered here the near-nadir component of backscatter. The ensounded area after this can be described as an expanding annulus (Figure 4). Returned acoustic signals can be broken down into near-nadir and off-nadir backscatter components in reference to the angle of incidence. The angle of incidence is the angle between the incident sound beam and a vertical line at right angle (i.e. nadir)

with the intercepting surface. Near-nadir backscatter is where the dominant response peak occurs and the backscatter intensity dissipates with the off-nadir signal. The peak amplitude response associated with the near-nadir backscatter component typically occurs within 20° off-nadir (Anderson, 2007). Eventually, the acoustic beam spreads out over the seabed to a point where there is little backscatter and the data is no longer representative of the seabed, at which point the data is said to reach its critical angle (Figure 5). In the context of this research, the critical angle is therefore defined as the angle at which the signal no longer contains relevant information on the seabed.

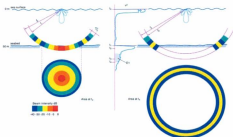


Figure 4: Sonarified area of a single acoustic beam as a function of time from a circle (left side) to an annulus (right side) on a flat horizontal seabed (source: Kloser, 2007).

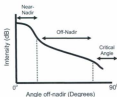


Figure 5: Acoustic energy return of a SBES footprint and how it can be used to measure the near-nadir and off-nadir zones (adapted from Anderson, 2007).

This project uses only the near-nadir component of the backscatter signal. This was done for a number of reasons. First, backscatter strength is dependent upon physical properties of the seabed, acoustic frequency and incidence angle. Standardizing the backscatter to the near-nadir to reflect the dominant response peak eliminates the influence of incidence angle, therefore leaving the backscatter response to be a function of seabed substrate properties and frequency. Second, the primary acoustic backscatter response is the "amplitude" of sound reflected from the seabed and this is measured most accurately from near-nadir backscatter response from the seabed. Third, near-nadir backscatter experiences less attenuation and therefore a higher signal-to-noise ratio (SNR). As a result, it may be more difficult to detect frequency differences in off-nadir backscatter. Fourth, Courtney et al. (2005) had found little difference in near-nadir and off-nadir metrics for defining sand and gravel classes on the Scotian Shelf, and the intensity of the signal (best measured near-nadir) as opposed to signal shape was the primary means of distinguishing substrate type. Therefore, the off-nadir component will not be used for further defining substrate types. There are certain seabed backscattering models that suggest that the near-nadir backscatter component is not as

effective as the off-nadir for discrimination seabed types (e.g. APL94, 1994). Even though near nadir data is used, the angles used to define near-nadir, (3° to 7° , see section 3.3 for further information) still include a range of backscatter differences among seabed types large enough to provide discrimination.

2.4 Acoustic Frequency

Sound is essentially a pressure wave, and the acoustic frequency is a measure of the number of wavelengths of a sound wave that pass through a point in a certain amount of time, measured in hertz (Hz). A high frequency would have more wavelengths per unit time as compared to a low frequency (Figure 6). The wavelength size (λ) is calculated by dividing the speed of sound (c) by the frequency (f) ($\lambda = c/f$).

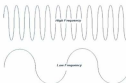


Figure 6: Low and high acoustic frequency waveforms (source: National Oceanic and Atmospheric Administration, NOAA).¹

Backscatter variations, as a function of frequency, are determined by the amplitude and wavelength of each frequency. These factors determine the ability of a sound wave to penetrate the seabed as a function of particle size. As the dimensions of a target (i.e. grain size) decrease so does the amount of backscatter (Lurton, 2002).

¹ http://www.wrh.noaa.gov/sewif/FAQs/EG-AFD_FAQs.php

Lower frequencies require relatively larger particles in order to reflect sound waves compared to higher frequencies. Therefore, the range of lower frequencies (10 kHz to 100 kHz) resolves less than higher frequencies but exhibits less attenuation in the water column, providing a larger maximum usable range (Lurton, 2002). This lower range of frequencies generally penetrates tens of centimetres into the seabed (Galloway and Collins, 1998), and carries more information back to the receiver due to this penetration that results in increased ensonification of the substrate's volume, and thereby including more substrate information (Collins and Rhynas, 1998). Hence, volume scattering and heterogeneities of the sediments are more important for lower frequency scattering responses (Holliday, 2007).

Higher frequencies (greater than 100 kHz) typically resolve more features due to its smaller wavelength providing greater resolution, but suffer greater attenuation in the water column limiting the maximum usable range (Lurton, 2002). Higher frequencies penetrate only centimetres into the seabed (Galloway and Collins, 1998). This is because sound waves from higher acoustic frequencies reflect off smaller particles and microscale surface roughness. Hence, surface scattering is more influential on high frequency backscatter values. High frequency backscatter is therefore representative of the surface veneer and smaller details of particles and roughness. Kloser et al. (2001) concluded, based on correlation coefficients from empirical studies, that a frequency of 120 kHz was most responsive to sediment size, as compared to 12 kHz and 38 kHz. Seabed roughness significantly influences acoustic backscatter if its dimensions are comparable in magnitude with the wavelength of sound (Lurton, 2002). As a rule of thumb, seabed roughness must be at least half the acoustic wavelength to significantly affect backscatter intensity at perpendicular angles (Femini and Flood, 2006). Several

studies have attributed backscatter from several frequencies to both surface and volume scattering (see discussion of these studies in section 1.2).

However, the relationship between backscatter strength and frequency at near-nadir angles is not always linear. One could reasonably assume that higher frequencies would produce higher near-nadir backscatter intensities simply because higher frequencies are more correlated with grain size and surface scattering causing the "peak response" of the signal to be greater. However, the model used in the APL94 (1994) report indicated that lower frequencies can have higher backscatter near-nadir than higher frequencies for certain substrate types. Certain empirical studies have also attained similar results (Chakraborty et al., 2007). Roughness could be the explaining factor here as higher frequencies at near-nadir may cause the signal to scatter away from the sensor due to its smaller wavelength. This would result in a decrease in the intensity of the signal while elongating the signal and shape of the return. On the other hand, Chakraborty et al. (2007) reported that lower frequencies exhibited more scattering from within the volume of substrate, and if the substrate type varies within the seabed from that of the surface than backscatter responses could vary. Therefore, there are a number of factors to consider when it comes to the expectations of backscatter returns from different frequencies and the eventual result depends, in some instances, on the parameters used in the model (see section 1.2 for a discussion of a number empirical studies on multi-frequency backscatter).

3. Methods

This section describes the study areas used in this project, the data collection process, the data processing and conditioning stage, and finally outlines the methods used for data analyses. Figure 7 provides a detailed outline of the methods used in this project. The process began with the collection and processing of raw acoustic backscatter data. Analyses were divided into two main components: univariate analysis of backscatter values from each frequency independently, and multivariate analysis techniques of the backscatter from both frequencies together.

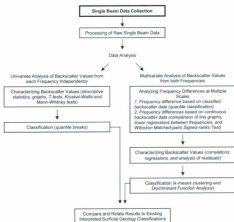


Figure 7: Detailed flowchart of the research method. Bold text indicates previous work not done by the author.

Univariate analyses involved both statistical analysis and classification of single frequency backscatter. Descriptive statistics and graphing of single frequency backscatter were conducted as an exploratory phase. The acoustic backscatter values were also analyzed using inferential statistics to test whether there were significant differences in mean backscatter values between study areas and seabed types. Then backscatter values were classified from each frequency independently using a quantile

classification method. Classification results were compared to the interpreted sediment units (ISUs) in order to assess the performance of univariate classification schemes.

Multivariate analyses involved the analysis of frequency differences, multivariate statistical analysis, and multivariate classification. Frequency differences were examined by directly comparing continuous backscatter values and classified backscatter values from each frequency. Multivariate statistical analysis techniques were used to understand the statistical relationship between backscatter values from two frequencies. Both supervised and unsupervised multivariate classification methods were used to delineate classes based on similar backscatter values and additional morphology variables. These results were again related to the ISUs.

3.1 Study Area

This research extends on previous work conducted by the Department of Fisheries and Oceans Canada (DFO). A funded research program was established on fish habitats in the offshore waters of the Scotian Shelf from 2002 to 2005. The Scotian Shelf is in the north-west Atlantic Ocean, just south-east of the Canadian province of Nova Scotia (Figure 8). The Scotian Shelf is a large continental shelf and an important area for the Canadian fishery as it has numerous fishing banks (Anderson et al., 2005). This DFO research program consisted of six study areas on three different fishing banks (Emerald, Sable Island, and Western). Two of these study areas on Western Bank were used for this study (Figure 8). Western Bank is a continuation of Sable Island bank and is an area that is representative of typical seabed environments on Canada's eastern continental shelf, being primarily composed of sand and gravel banks (Courtney et al., 2005). The study areas are characterized as being relatively simple in terms of particle size distributions on the Wentworth scale and extremely flat in terms of seabed slope.

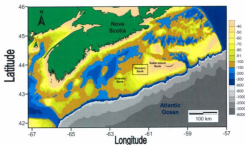


Figure 8: Study area displaying Western Bank (rectangle) on the outer Scotian Shelf. Banks are represented by areas shallower than 91 m depth (50 fathoms). Shaded contours represent depths in meters.

A previous study conducted by Anderson et al. (2005) had identified two 100 km² study areas on Western bank based on preferred and non-preferred juvenile Haddock habitat (Figure 9). The two study areas were approximately 15 km apart. The depth approximately ranged from 49 m to 60 m in the 'preferred' study area and from 47 m to 55 m in the 'non-preferred' study area. In addition, there were two smaller 5 km² detailed study areas within each of the larger study areas. Detailed study areas were chosen based on high measures of bathymetric relief in comparison to the rest of the 100 km² study area. The preferred study area had a greater range of depths, which may indicate greater complexity of habitats (~11 m as compared to ~9 m in the non-preferred study area) and bathymetric relief. This is expected as juvenile Haddock seek benthic habitats that are more rough and complex to provide protection from predation (Anderson et al.,

2005). Both the 5 km² preferred and non-preferred detailed study areas were used as the study areas for this study.

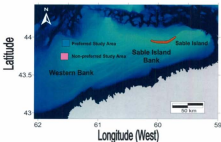


Figure 9: Western Bank displaying the location of the two study areas, Preferred (upper box) and non-preferred (lower box) juvenile Haddock habitat areas. The shaded bathymetric relief ranges from shallow (light blue) to deep (dark blue) where depth greater than 1000 m off the shelf are shaded gray.

The selection of this particular study area was done for a number of reasons. First of all, there has been a lot of work done in this area (Anderson et al., 2005, Courtney et al., 2005, Anderson et al., 2007), particularly by DFO, on acoustic seabed mapping as well as habitat mapping and management. These studies provided some useful information on the marine environment of Western Bank, including seabed geomorphology characteristics and fish habitat distribution. Secondly, this is one of the few areas where two single beam acoustic frequencies were collected and made available by DFO. Previous work with these SBES data, as well as sidescan, multibeam, and geological ground-truthing, has included both geological and biological habitat

mapping (Anderson et al., 2005, Courtney et al., 2005, Anderson et al., 2007, Fader, 2007, Ollerhead and Anderson, 2007). The area has little flora present which could otherwise interfere with acoustic echoes from the seabed (Fader, 2007).

An important asset for this project is that there exists, for these study areas, maps of surficial geology interpreted by marine geologists (Figure 10 and Figure 12), using a combination of sidescan imagery and ground-truthing using physical samples (Fader, 2007). The sidescan imagery provided a mosaic that highlights the shape, texture, and roughness of the seabed at a 0.25 m resolution. Sidescan images allow for visual interpretations of roughness and texture because data are restricted to the wide-angle, off-nadir response (Anderson, 2007). Marine geologists interpret sidescan data by analyzing them as an image of the seabed portraying the texture and composition, as well as examining the relative backscatter of reflected energy. Sediment particle size analyses and photographic records provide verification or ground truthing for their interpretations of the image and provide information on the texture of sediments and particle shape (Fader, 2007). Geological interpretations of the study sites produced fourteen geological classes, called interpreted sediment units (ISUs), based on morphology, grain size, and bedforms (Fader, 2007). This information is used in this study as ground-truthing, being compared to classifications obtained from the acoustic data. There is an obvious bimodal distribution of depth in the non-preferred study area (Figure 11) where the southwestern section is deeper (>52 m) than the northeastern section (<52 m). In the non-preferred study area (Figure 10), the southwestern portion with depths >52 m consists only of five ISUs (GH, GL, GRL, GRS, and SB) and the northeastern portion with depths <52 m consists of the remaining three ISUs (GRI, SM, and Sand).

Interpreted surficial geology of the non-preferred study area

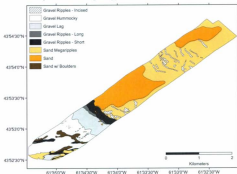


Figure 10: Surficial geology of the non-preferred study area. These data were produced by marine geologists at NRCan (Fader, 2007).

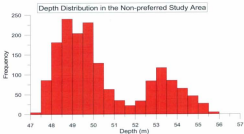


Figure 11: Histogram displaying the bimodal depth distribution in the non-preferred study area.

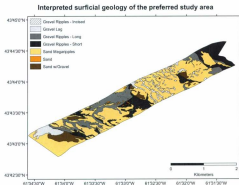


Figure 12: Surficial geology of the preferred study area. These data were produced by marine geologists at NRCan (Fader, 2007).

The current geological setting of the bank has been shaped by events following the last glaciation and has been previously described by Courtney et al. (2005) and Fader (2007). Marine transgressions have modified pre-existing glacial deposits in the area into well-sorted sands and gravels, while fine-grained silts and clays have been transported into deeper water (Courtney et al., 2005). Sediments are generally reworked and transported from the northeast to southwest by waves and currents (Courtney et al., 2005). Recent acoustic surveys have revealed that the seabed contains evidence of glacial processes, as evident by moraines, and fluvial erosion and there is a high degree of sediment patchiness and small-scale roughness (Fader, 2007). There is a full range of sediment bedforms, which include ripples, megaripples, sand ribbons, and ridges (Fader,

2007). In sum, the seabed in this study area can be portrayed as a mobile sand layer overlying a relict gravel lag deposit (Courtney et al., 2005). Each ISU requires its own explanation and is explained in Table 1.

Table 1: Description of the Interpreted Sediment Units (ISUs).

Name	Description	Reference
Gravel Lag (GL)	A layer of coarse gravel caused by the removal of finer sediments and is therefore less susceptible to movement by waves and currents. Gravel lag appears as a flat surface with uniform backscatter.	Courtney et al., 2005
Gravel Hummocky (GH)	GH is essentially a combination of gravel lag and boulders and is "... a rough and hummocky surface at the meter scale and ridges of boulders are widespread".	Courtney et al., 2005, p.3
Gravel Ripples Short (GRS) and Long GRs Wavelength	Both are similar texturally and are sloping surfaces called ripples, or waves, of loosely compacted sediments composed of granules, pebbles and cobbles. They are interpreted to be a result of wave energy reaching the seabed during large storms. They have formed under slightly differing energy conditions and are separated by wavelength size, short wavelength (2 to 3 m) and long wavelength (> 5 m).	G. Fader, Personal Communication, November, 26 th , 2009 Fader, 2007 Courtney et al., 2005
Gravel Ripples Incised (GRI)	GRI are incised unilinear depressions of sandy gravel within the SM class and are likely the troughs of large sand bedforms. The production of coarser sediments results from the winnowing of sand from the seabed and then gets formed into gravel ripples.	Courtney et al., 2005 G. Fader, Personal Communication, November, 26 th , 2009
Sand (Sand) and Sand with Megaripples (SM)	These are distinguished by flat areas versus areas with ripples. Megaripples typically range in height from 0.1 – 2.0 m.	Ames and King, 1984
Sand with Boulders (SB)	Are dispersed throughout the GR class and is more of a mixed sand and gravel class with boulders.	Courtney et al., 2005
Sand to Sandy Gravel (SG)	A mixture of sand and gravel	

The spatial arrangement and area of these ISUs vary between study areas (Table 2). The ISUs that constitute seabeds with dominant sand substrate include SM, SB, SG, and Sand, and the ISUs for the seabeds with dominant gravel substrate are GRS, GRI, GRl, GH, and GL. Sand seabeds dominate gravel seabeds in each study area with the non-preferred study displaying a higher percentage of sand seabeds as compared to the preferred study area. The average area of gravel seabeds is larger than the average area of sand seabeds in both study areas, meaning that sand seabeds are

smaller and less continuous (i.e. more dispersed) than gravel seabeds. The preferred study area has a much more complex spatial arrangement of ISUs and has a larger number of smaller ISUs as compared to the non-preferred study area.

Table 2: Summary table of sand and gravel polygons on Western Bank.

Study Area	Total # of polygons	Mean area per polygon (m ²)	# of ISUs	% of sand seabeds	% of gravel seabeds	Total area (m ²)	Mean area per sand polygon (m ²)	Mean area per gravel polygon (m ²)
Non-preferred	163	117,131	6	96.9	34.9	4,540,218.99	110,362	123,679
Preferred	206	32,079	7	59.7	40.3	4,014,304.10	25,391	38,767

The mean area and percentage coverage of each ISU also varies within and between the study areas (Table 3). SM, followed by Sand and GH, predominantly cover the non-preferred study area. SM covers the largest percentage of the preferred study area, followed by GRS and GRL. Therefore, SM is a dominant seabed type in both study areas, more so in the preferred study area. As indicated by the mean area per ISU, GH polygons are very large, and to a lesser extent, so are the SM polygons in the non-preferred study area. In the preferred study area, the largest polygons are SM followed by GL and GRL.

Table 3: Summary table of the mean area (m) and percent (%) of coverage of each ISU in comparison to the whole study area.

ISU	Non-preferred Study Area Depth > 52m - Mean Area per ISU (x10 ⁶ m ²)	%	Non-preferred Study Area Depth < 52m - Mean Area per ISU (x10 ⁶ m ²)	%	Preferred Study Area - Mean Area per ISU (x10 ⁶ m ²)	%
GRS	45.4	5.0	-	-	38.8	17.3
GRL	12.0	9.8	-	-	54.5	14.9
GRI	-	-	6.8	3.7	13.2	4.9
GL	117.7	7.8	-	-	63.6	3.2
GH	756.7	16.7	-	-	-	-
SG	-	-	-	-	11.3	6.7
SM	-	-	212.8	37.5	70.5	52.6
Sand	-	-	21.3	26.3	0.5	0.4
SB	21.4	2.4	-	-	-	-

Attempts to statistically classify these ISUs based on multibeam (455 kHz), sidescan (120 kHz), and normal incidence echo sounders (120 kHz) demonstrated that acoustic backscatter from the seabed occurred in four unique statistical classes (Courtney et al., 2005, Ollerhead and Anderson, 2007). These classes were attributed to sand and gravel, but further refinement of the classes and attempts to relate to the ISUs remain ambiguous. Attempts to relate the spatial structure of the four acoustic classes to interpreted surficial geology features succeeded at broad spatial scales, but were unsuccessful at smaller spatial scales across 10s to 100s of meters. It is anticipated that a dual-frequency ASC approach can improve upon these efforts.

3.2 Data Collection

Data collection was not conducted by the author but rather by DFO research scientists. Extensive acoustic datasets were collected in 2002 and 2003, as part of the DFO research program discussed in section 3.1, in both the 5 km² preferred and non-preferred habitat study areas of Western Bank. Fully calibrated backscatter data were collected using normal incidence SBES acoustic systems (different systems for the 2002 and 2003 surveys), running at 38 kHz and 120 kHz frequencies. A BioSonics DT6000 calibrated scientific echo sounder was used for the 2002 survey and a BioSonics DT-X calibrated scientific echo sounder was used for the 2003 survey. Calibrations of the two echo sounders were performed by technicians from Fisheries and Ocean Canada using standard calibration procedures from Foote et al. (1987). Data from both frequencies were collected simultaneously from surveys done during research expeditions aboard the Canadian Coast Guard Ship (CCGS) *Hudson*. Vessel speed was held constant at approximately 1.54m/s for all surveys. In 2002, data were collected from September 21st to October 2nd and in 2003 from October 4th to October 14th. In 2002 transect lines 2 to 5

were collected and in 2003 lines 7 and 8 were collected (Figure 13), in the non-preferred study area, survey lines 2, 3, 7, and 8 were surveyed from northeast to southwest, whereas survey lines 4 and 5 were surveyed from southwest to northeast. The same design was used in the preferred study area, except that survey line 7 was surveyed from southwest to northeast. Survey lines were 5 km in length and average spacing between survey lines was 200 m. There were also four additional lines (1a, 1b, 6a, 6b) collected in 2002 that lie between two outer lines on each side. These lines were surveyed with a Towcam (a camera towed close to the bottom that captures video) for a biological survey (Gordon et al., 2000), which collected data on the presence or absence of fish at day (lines 1a and 6a) and night (lines 1b and 6b). These lines are not displayed on the map in Figure 13 for the purposes of clarity.

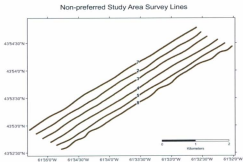


Figure 13: Acoustic survey lines 2 to 8 for the non-preferred study area.

The SBES systems ran both frequencies simultaneously using two transducers, one for each frequency. The data collected were time-stamped and geo-referenced using a differential GPS. In 2002, data were collected at a rate of 1 ping per second (pps) with the transmitter running at 2 pps, alternating between the two frequencies. In 2003, data were collected at a rate of 2 pps. The 38 kHz frequency is a split beam system with a beamwidth of 5.6 degrees and the 120 kHz frequency system uses a concentric dual beam configuration, narrow and wide, with 6.5 and 17.8 degree beamwidths, respectively. Both frequencies have a 0.4 ms pulse length. The transducers were fitted to a hydro-dynamic boom fixed to the starboard rear quarter of the CCGS Hudson, where the transducer heads were approximately 5 m below the sea surface. The analog-to-digital conversion featured a wide dynamic range of approximately 132 dB and a sampling rate of approximately 42 kHz. The raw echo data from the surveys were logged for post-processing in a compressed format known as BioSonics DT4.

In addition to the single beam data and interpreted surficial geology maps for the two study areas, there also exists a multibeam echo sounder (MBES) bathymetric dataset. A 100km² (10km x 10km) area is covered by the MBES survey in both study areas and provides detailed bathymetry at 8m resolution. These data were processed in ESRI's ArcGIS to provide additional geomorphology layers. These processes will be further explained in section 3.8.

3.3 Data Processing

The first step in this study was to take the BioSonics DT4 raw calibrated acoustic data and process them (Figure 14) for radiometric and geometric biases (i.e. remove the effects induced by the acoustic system and the transducer's varying range from the seabed). A number of factors other than physical seabed characteristics affect seabed

backscatter values. These factors include changes in depth, incidence angle, ship motion, and sound speed variations in the water column, as well as the configuration of the acoustic system, including its frequency, pulse length, beam angle, and beam pattern. However, in the context of this research it is important to minimize any affect on the acoustic signal other than those caused by physical seabed properties and the acoustic frequency. This helps to ensure that backscatter intensity values are dependent upon physical seabed characteristics alone thereby providing "clues" for seabed classification. Hence, it is important that "[t]hese classification "clues" will ... only [be] weakly dependent upon nonintrinsic factors such as propagation, pulse length, and transducer directivity and should, therefore, provide a basis for robust seabed classification algorithms" (Jackson et al., 1996, p. 458).

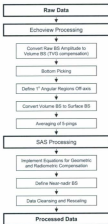


Figure 14: Diagram of the raw acoustic data processing steps using both Echoview and SAS computer software (BS = backscatter, TVG = Time-Varied Gain).

Only the 38 kHz datasets had to be processed for this research project as the 120 kHz datasets had been previously processed by DFO using the same method. The acoustic data were processed using third party software in order to extract surface backscatter (BS) data from the echo envelope. The acoustics processing software Echoview (Myriax Inc.) and Statistical Analysis Software (SAS, by SAS Institute Inc.) were used for this process. The process was developed by Chris Lang, a marine

acoustics scientist with the North-west Atlantic Fisheries Centre, Department of Fisheries and Oceans.

Echoview processing is dependant upon the echosounder's configuration (i.e. frequency, beamwidth, pulse length) and the environmental parameters (i.e. seawater propagation parameters such as sound speed and absorption coefficient based on temperature, salinity, and pressure measurements) at the time of the survey (Table 4). There is a slight difference in the beam angles between the 2002 and 2003 data, which is a result of changing the system from a DT6000 to a DT-X. This difference is small (0.2° , or 0.7° for the 120 kHz wide beam) and all data from both years were compensated and standardized to the near-nadir peak backscatter response signal using specific incidence angles. As a result, the small beam width difference between the years is minimized. Hence, the echosounder's configuration was constant in terms of beamwidth and pulse length for each year, leaving only frequency as the changing characteristic. In addition, it was assumed that the environmental parameters remained constant over the study area for the time of each survey. For this, a CTD (conductivity-temperature-depth) cast was done in each study area using a calibrated Seabird-19. The input and processing of the information in Table 4 helped to provide compensated echograms of depth and backscatter intensity. The 120 kHz wide beamwidth channel was used.

Table 4: Input parameters for data processing using Echoview.

Frequency	Year	Average Sound Speed	Beamwidth	Pulse Length	Absorption Coefficient
38 kHz	2002	1468.7 m/s	5.6°	0.4 ms	0.015 dB/m
38 kHz	2003	1465.1 m/s	5.8°	0.4 ms	0.010 dB/m
120 kHz	2002	1468.7 m/s	Narrow 6.5° Wide 17.8°	0.4 ms	0.0405 dB/m
120 kHz	2003	1465.1 m/s	Narrow 6.3° Wide 18.5°	0.4 ms	0.033 dB/m

Sound levels extend over many orders of magnitude and because of their huge dynamics, acoustic values like intensity or pressure of a sound wave, are typically quantified on a logarithmic scale and expressed in decibels (dB) (Lurton, 2002). "By using the decibel scale, calculations are simplified and relative values relate more closely to perception (Pike, 1996)." The decibel is calculated as 10 times the logarithm of the ratio of the intensity (I) level of a sound wave to a reference intensity level:

$$I \text{ (dB)} = 10 \text{ Log}_{10} (I_{\text{sound}} / I_{\text{reference}})$$

Intensity of a sound wave is the average amount of energy transmitted per unit time through a unit area in a specified direction. Power is the amount of energy transmitted per unit time and is measured in watt; therefore, intensity is measured in watts per square meter (DOSITS, 2009). However, underwater acoustic systems typically measure pressure instead of intensity, and intensity of a sound wave is proportional to the square of its pressure. The intensity in dB can be computed from the pressure:

$$I \text{ (dB)} = 10 \text{ Log}_{10} (p^2_{\text{sound}} / p^2_{\text{reference}}) = 10 \text{ Log}_{10} ((p_{\text{sound}}/p_{\text{reference}})^2) = 20 \text{ Log}_{10} (p_{\text{sound}} / p_{\text{reference}})$$

Hence, 10 Log is used for quantities of acoustic power and intensity and 20 Log is used for quantities of acoustic pressure (Lurton, 2002). In this research backscattering coefficients are used, which are ratios of intensities and hence 10 Log is used. A decibel value is a ratio and therefore a relative value with no units, in other words a dimensionless number. Therefore, a reference intensity is needed to give absolute intensity or pressure levels in decibels and is usually the microPascal (μPa) (Lurton, 2002).

$$p_{\text{reference}} = 1 \mu\text{Pa}$$

Because Echoview is primarily used to process data for fisheries acoustics applications, it normally provides data in two compensation (TVG – time-varied gain)

domains: 20 Log R, used for reporting volume backscattering strength (S_v), and 40 Log R, used for reporting point target strength (TS), where R refers to the range from the seabed (Echoview, 2010). Backscatter strength will be a function of the targets distance from the transducer and TVG compensates for this propagation loss. Volume backscatter data (S_v , or 20 Log R) (see Echoview (2010) for the formula used to convert and save raw BioSonics backscatter as S_v) was converted to surface area backscatter (SA, or 30 Log R) within Echoview by adding 10 Log R to each measurement (Chris Lang, DFO Scientist, Personal Communication). This is the most logical form of data representation as the seabed is an extended target and neither the volume nor the target strength domains are optimal.

Before the data were converted to the 30 Log R domain (surface area backscatter), the ocean bottom (i.e. seabed) was selected from each ping. For this process, Echoview provides a bottom-picking algorithm, called maximum backscatter with backstep. This technique searches each ping to find the data sample with maximum volume backscatter strength (which is usually caused by the seabed interface). It then starts from this sample and searches up the ping to find the first sample with a backscatter value less than a specified discrimination level (-30 dB was used for this study). Following bottom detection, thirty-four 1-degree angular regions were defined for each ping relative to the detected bottom. Data were then converted to the 30 Log R domain and binned into five-ping intervals for each angular region. Averaging over five ping intervals reduces ping-to-ping variability caused by stochastic noise (i.e. this increases the signal-to-noise ratio, SNR) and also minimizes measurement error associated with seabed slope, ship motion, and differential GPS positioning (± 3 m). Therefore, each observation is an average across all measurements in an angular region over five pings. The latitude and longitude coordinates for each 5-ping sample was

calculated as the average of the five pings. Komelussen and Ona (2002) suggest it is useful to average several pings to reduce variability expected in acoustic data, and argue that it provides better discrimination between acoustic targets. The datasets were exported from Echoview for use in the SAS software.

The SAS processing further compensated the data for beam pattern effects considering varying ranges and incidence angles. Two equations were implemented in SAS in order to compensate the backscatter values for changes in footprint size and incidence angle of each angular region as a result of depth variations. This compensation is required because the final output measures backscatter per unit area and it is important that each angular region is representative of the same unit area. These processing steps also compensated for changes with respect to where each angular region is measured along the beam pattern as a result of changing depth. The greater the angle of incidence of a specific angular region along the beam pattern, the weaker the acoustic energy return becomes. Therefore, changing depth not only influences backscatter values for specific angular regions through changing footprint size, but also by changing the location along the beam pattern where each region is measured. The equations are based on the assumptions of no ship movement and a flat seabed (i.e. no slope). These assumptions will be sources of noise in the backscatter data, and the five-ping averaging in Echoview reduces these noise effects. However, sampling did not occur if the sea state was rough and it is assumed that the variability in sea states is a small order effect (John Anderson, DFO Scientist, Personal Communication). The first equation calculated the incidence angle for each of the angular regions of the 5-ping intervals. The incidence angle of any particular Echoview angular region will vary based on changes in depth (bigger angles at smaller depths). Therefore, the equation calculated a middle incidence angle for each angular region of

the five ping intervals using a region-specific and transducer-specific equation. This equation has built-in assumptions that Echoview detected the seabed while the pulse length of the acoustic echo was so far into the seabed. More specifically, the pulse length of the acoustic systems used in the surveys was 0.4 ms; therefore, an acoustic beam is not fully ensonifying the seabed until 0.4 ms from the initial contact. The equations used here to calculate incidence angle assume that the acoustic echo has detected "bottom" when half the pulse length (0.2 ms) is ensonifying the seabed. The generic form of the incidence angle is:

$$\text{Angle}(H) = b_{0i} + b_{1i} \cdot H + b_{2i} \cdot H^2$$

where H is the transducer height above the seabed, the b values are three region and transducer-specific coefficients, and the subscript, i , represents each Echoview region. For example, the equation for angular region 7 (as defined in Echoview) for one 120 kHz dataset is:

$$\text{Angle}(H) = 10.5724 - 0.05006 \cdot H + 0.000261 \cdot H^2$$

and if, over a whole survey line, H varied from 43.3 m to 50.6 m, then the incidence angle (of angular region 7 backscatter) would vary from 8.89 to 8.71 degrees.

The incidence angle information is then used in a second equation which compensated, or normalized, the surface area (30 Log R) backscatter values for changes as a result of the following factors: 1) varying footprint sizes due to depth changes, 2) incidence angle differences due to depth changes, 3) and beam pattern variations. It is necessary that each angular region has surface area backscatter values normalized (per unit area) along the whole survey line, which allows them to be compared. The equation compensated for changing footprint sizes by using the depth values. The incidence angle is also compensated for, as changing depth determines incidence angles, which also influences backscatter value's area compensation. Finally,

this equation also considered where, along the beam pattern, a specific region is measured. This is important because the further from the acoustic axis a region is measured, the weaker the signal gets. The following equation was used for compensation of each region's changing footprint size, incidence angle, and beam pattern:

$$BS_i(Out30,H) = Out30_i + a_{0i} + a_{1i} \cdot H + a_{2i} \cdot H^2$$

$Out30_i$ represents the Echoview backscatter output for a specific region, the a values are the coefficients, and H is the transducer's height above the seabed. The result of this processing produced calibrated and compensated surface area backscattering strength values for a number of one-degree increment incidence angles from two frequencies.

The data were further processed to derive the near-nadir backscatter component to emphasize amplitude differences resulting from changing the acoustic frequency. The near-nadir component of backscatter can be defined by examining graphs of average backscatter intensity for all 34 angular regions for each survey line. A typical graph appears as a series of 'humps' representing the nature of the echo returns (Figure 15 and Figure 16). The first 'hump' is the initial dominant peak return of acoustic energy from the seabed to the system, and is representative of the near-nadir acoustic backscatter. Each survey line was examined graphically for angular backscatter intensity values, and angular regions representing this dominant peak response near-nadir were defined for each frequency. The angular regions determined for the near-nadir response ranged from 3° to 6° for the 38 kHz frequency and from 4° to 7° for the 120 kHz frequency. The difference in incidence angles for the near-nadir component between frequencies is small and these angles were chosen to standardize the near-nadir backscatter to dominant response peaks. Therefore, the difference in incidence angles between the two frequencies is not considered as being a major issue. The difference in

the angular regions of the near-nadir component is a result of the slight differences in how the 38 kHz and 120 kHz systems were built and their different parameters (see Table 4). These defined angular regions were used to produce four new datasets (two frequencies for each study area). These datasets have near-nadir backscatter values for each point (i.e. 5-ping interval) which are averaged backscatter values for the specific near-nadir angular regions.

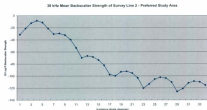


Figure 15: Line graph displaying an example of average backscatter strength values (dB) from the 38 kHz frequency as a function of incidence angle (degrees). Near-nadir response is defined from 3° to 6°.

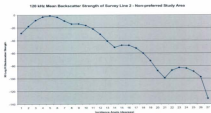


Figure 16: Line graph displaying an example of average backscatter strength values (dB) from the 120 kHz frequency as a function of incidence angle (degrees). Near-nadir response is defined from 4° to 7° . X-axis is cut-off at 27 degrees due to extremely low off-nadir values.

Once the near-nadir datasets were produced it was discovered that some values were out of normal range. There were positive backscatter values for the 120 kHz data and this was unexpected as backscattered echoes from the seabed are typically lower than the initial acoustic signal due to the amount of absorption and scattering the signal experiences. Therefore, all of the 120 kHz backscatter data had to undergo an 8 dB adjustment (8 dB was the highest value above 0 dB in the datasets) downwards (i.e. subtraction) to ensure that all values were rescaled below 0 dB. In addition, the 2003 backscatter datasets were rescaled to compensate for the re-configuration of the BioSonics DT hydroacoustic system from a DT6000 to a DT-X. While the transducers remained the same, the electronics associated with the digital component were significantly changed by this upgrade, which was equivalent to purchasing a new echo sounder. Therefore, adjustments were necessary to standardize the acoustic data between the two systems. The 38 kHz datasets had 3.0 dB subtracted and the 120 kHz had 3.7 dB subtracted in order to make the 2003 backscatter values equivalent to the

2002 backscatter values. The values were calculated as the mean difference between backscatter values from each year for each frequency. These adjustments were made to ensure the data were statistically equivalent based on the offsets applied. Finally, several points along each survey line for the 38 kHz frequency backscatter values were way out of normal range due to glitches in the acoustic system. These anomalies were obvious in the echograms and caused by erroneous pings. These points were filtered out and removed from the datasets. These removed points accounted for 2.7% (167/5966) of all the data points in the preferred study area and 2% (111/5368) of the data in the non-preferred study area. Removal of such values is a necessary first step to avoid statistical biases.

An additional issue with the data was discovered following the aforementioned data processing steps. It was assumed that each survey line was starting with a 120 kHz data point and ending with a 38 kHz point and that both frequencies alternate. Therefore, each survey line should have the exact same number of points for each frequency. However, on average there are about 4 or 5 extra points per survey line for one of the two frequencies. This is a fault of the acoustic system and in order to compare the two frequencies, the datasets should have the same number of points and match spatially. Any points not matching the assumptions were removed. For example, if a survey started with a point from the 38 kHz frequency, then this point was removed. To allow some statistical analysis, the geographical coordinates of each set of points (a 38 kHz point and a 120 kHz point) were then averaged to make them overlap. Note that the points in these datasets did not exactly overlap between frequencies, as there was a slight delay between pings from each frequency. The horizontal geodetic datum used for the geographic latitude and longitude coordinates was the World Geodetic System of 1984 (WGS 84). The overall shift (1 m – 5 m) applied to each set of points remained well

inside the footprints of the two original samples so that the new location was still representative of the original sampled area. These datasets, for which the location of the 38 kHz and 120 kHz samples were averaged, are referred to as the balanced datasets (Appendix G) and are used when the two frequencies are directly compared.

The size of the footprint per unit area is calculated based on the angle chosen to define the near-nadir backscatter component, which acts as the beam angle, and the water depth at a specific sampled site. The footprint is the amount of area that the acoustic beam encompasses on the seabed. Because five pings were binned, the footprint of one data point actually takes a rectangular shape with curved ends. Therefore, the width of the footprint is the diameter and the along-track length of the footprint is calculated using the equation (diameter + (vessel speed * 4)). The dimensions of footprint size for the 38 kHz frequency ranged from 8.5m (width) x 14.5m (along-track length) to 12.6m x 18.8m, and the 120 kHz ranged from 9.8m x 16m to 14.7m x 20.9m. The larger footprint size for the 120 kHz frequency is due to the larger angle chosen to define the near-nadir component. The average distance between each 5-ping sample was 12 m, meaning that the overlap between the footprints of each sample could almost reach up to 50%.

The speed of sound was calculated during the survey using measures of temperature, salinity, and pressure. The average speed of sound between both years was approximately $1465 \text{ m} \cdot \text{s}^{-1}$. Therefore, the wavelength size for the 38 kHz frequency is 3.85 cm and for the 120 kHz frequency the wavelength is 1.22 cm. This is calculated based on the following formula:

$$\lambda = c/f$$

c is the speed of sound, and f is the frequency (Hz).

3.4 Univariate Acoustic Backscatter Analysis

Statistical results throughout this manuscript are based on decibels, which are logarithmic values (i.e. the exponent) of backscatter intensities and, therefore, formally have no units; they are a dimensionless number that satisfy the assumptions of normal distribution. Hence, mean values represent the geometric mean, not the arithmetic mean, with no units associated with results. These logarithmic values still represent the original backscatter intensity data.

For each acoustic frequency, near-nadir backscatter values (dB) from both study areas (i.e. using all survey lines within a study area) were plotted using histograms. The data consisted of the eight survey lines of data collected in 2002 (1a, 1b, 2, 3, 4, 5, 6a, 6b) and the two lines from the 2003 survey (7 and 8) for each study area. The original backscatter datasets (i.e. datasets that did not have each set of dual-frequency observations spatially matched) were used for all univariate analysis.

Descriptive statistics (i.e. mean, variance, standard deviation, skewness, and kurtosis) provided a quantitative assessment of backscatter data. The coefficient of variation is a relative measure of statistical variation (McGrew and Monroe, 2000) and was a useful measure in this context as two completely different study areas and two different acoustic frequencies were compared and it was expected that the magnitude of backscatter intensity would differ based on these parameters. The significance test for skewness is taken from Tabachnick and Fidell (2001). The standard error for skewness (S_s) is expressed by taking the square root of six divided by the number of cases (N):

$$S_s = \sqrt{6 / N}$$

This skewness value (S) is then compared with zero using the z distribution:

$$z = (S - 0) / S_s$$

Like typical hypothesis testing, if the skewness value is outside of the z -value then the skewness value is significant. Kurtosis values indicate peakedness in the data, which in a unimodal distribution can be used as a proxy for clustering.

Independent samples T-tests were used to determine whether mean backscatter values were significantly different between study areas for each frequency. The Levene's test for equality of variances was used to determine whether the corresponding variance values were homogeneous.

Descriptive statistics of backscatter values from each frequency were also summarized by combining dominant sand substrate seabeds (i.e. Sand, SM, SG, and SB) and dominant gravel substrate seabeds (i.e. GRS, GRL, GRI, GL, GH). Independent samples T-tests were used to determine whether mean backscatter values were significantly different between sand and gravel seabeds for each frequency in both study areas. Descriptive statistics were also calculated for each ISU from the interpreted surficial geology maps. Only four of the survey lines (lines 2, 3, 4, 5) from each study area actually overlap the interpreted surficial geology maps. Therefore, these survey lines were the only ones used for analysis of specific seabed types.

The final step in the univariate analysis was to statistically test mean backscatter values for significant differences between ISUs for each acoustic frequency. The ANOVA three-or-more sample difference F test would be the appropriate test to be used here, however, the Levene statistic, which tests the assumption of equal variances between groups (SPSS, 2007), revealed that the variances between the ISUs were not equal and therefore the assumptions of ANOVA were broken. A non-parametric Kruskal-Wallis H test was hence deemed appropriate. The Kruskal-Wallis H tests three or more sample means for significant differences without the assumptions that the variances of each sample are equal (McGrew and Monroe, 2000).

The Kruskal-Wallis test did not reveal which ISUs were significantly different from every other ISU. The Wilcoxon Rank Sum W and the Mann-Whitney U test are the non-parametric statistical tests used for determining whether one ISU was significantly different from another in terms of ranked values of backscatter values from a particular frequency. The Wilcoxon rank sum W and the Mann-Whitney U tests uses the sum of ranked positions from a particular variable from two different sets of sampled data to measure the magnitude of their differences (McGrew and Monroe, 2000). More specifically, values from two different samples are combined and ranked, then the sum of the ranked values for each independent samples (i.e. geological class) are tested for significant differences. The test statistic (Z_w) for the two-sample Wilcoxon rank sum W test simply subtracts the mean rank (\overline{W}) from sum of the ranks (W) and divides it by the standard deviation (s_w), all for the sample with the smaller sample size:

$$Z_w = (W - \overline{W}) / s_w$$

The Wilcoxon rank sum W and Mann-Whitney U tests are very similar, and the test statistic values are equivalent in the sense that they both provide the same significance level (McGrew and Monroe, 2000). The null hypothesis states that the distribution of measurements for the first population is equal to that of the second population. This is a two-tailed hypothesis, as there is no assumption made about the direction of the relationship between any pairs of the geological classes.

3.5 Univariate Acoustic Backscatter Classification

Before classifying the seabed using a number of input variables in a multivariate classification, it was first assessed how well classified backscatter values from a single frequency could represent seabed substrate types on the Scotian Shelf. The balanced datasets of backscatter values (spatially matched 38 kHz and 120 kHz samples) from

each frequency were classified using a quantile breaks classification method. This method creates classes that have an equal number of observations. This is a way to normalize the data and compare backscatter between frequencies on a relative scale. Results were overlaid on the interpreted surficial geology maps using GIS software for comparison with seabed types. In addition, the number of observations for each quantile class was assessed for seabeds with dominant sand and gravel substrate types, as well as each ISU.

3.6 Frequency Differences

The results of the quantile classification of backscatter values of each frequency were differenced to assess relative frequency differences based on classified backscatter values. The following equation was used:

$$\text{Quantile Class Difference} = 120 \text{ kHz Quantile Class} - 38 \text{ kHz Quantile Class}$$

The number of observations for each difference value was summarized by study area, by sand seabed types (combined ISUs with dominant sand substrate) and gravel seabed types (combined ISUs with dominant gravel substrate), and by each ISU. The differencing results were also overlaid onto the interpreted surficial geology map. In addition, results of the quantile classification differencing were classified as no difference, higher quantile class for the 38 kHz frequency, or higher quantile class for the 120 kHz frequency. A measure of spatial continuity (i.e. distance or spatial scale) of these relative frequency differences was calculated by measuring the number of occurrences of adjacent points with the same value (i.e. no difference in quantile classes, higher quantile class for 38 kHz, and higher quantile class for 120 kHz). The dominant scale at which continuous points occurred for these three values was analyzed.

Examining the divergences between backscatter values from the two frequencies was also done using continuous (i.e. unclassified) backscatter values. Backscatter values from both frequencies of the balanced datasets were smoothed using a running average, in order to reduce variability between adjacent observations. The smoothed 9-point backscatter values from both frequencies were plotted on line graphs to examine the general trends and the spatial scales at which these trends occurred along each survey line. These smoothed backscatter data from both frequencies were then directly compared to identify frequency divergences within certain window sizes that contained a specific number of samples. To analyze frequency divergences at different spatial scales, four different sized moving windows were used and moved along all the observations in order to compare smoothed backscatter values from both frequencies. The window size started with five observations and increased each window size by ten (5, 15, 25, and 35). Within each window, the slope of the linear regression line for each frequency was directly compared, and if the slope values were different, (i.e. positive vs. negative, or vice versa) the center point of that window was flagged as having a frequency divergence. The points with divergences were statistically summarized and plotted on the surficial geology map to highlight places where 38 kHz and 120 kHz diverged.

Finally, a matched-pairs (dependent-sample) difference test was used to determine whether continuous backscatter values from both frequencies were significantly different for each ISU. For this test, there were two sampled datasets (i.e. balanced backscatter datasets from each frequency) compared to each other for one group (i.e. ISU). As a result, the most appropriate statistical analysis for this comparison was the non-parametric Wilcoxon Matched-pairs Signed-ranks Test (McGrew and Monroe, 2000). A non-parametric test was used because the distribution of backscatter

values for certain ISUs were not normal. This test differences the values of each matched-pair and then ranks these differences. The difference could be either positive or negative. The positive difference ranked values are summed and the negative ranked values are summed. The test used only one of the two possible rank sum values. This is a non-directional two-tailed test because no direction of difference between the variables is hypothesized, and therefore the smaller of the positive and negative rank sum values is used (McGrew and Monroe, 2000). The backscatter values from both frequencies were normalized to a [0,1] scale for this test. This was achieved using the following equation:

$$\text{Normalized Value} = (\text{original value} - \min(\text{all values})) / (\max(\text{all values}) - \min(\text{all values}))$$

Backscatter data from the two frequencies needed to be normalized so that backscatter values could be directly compared. This technique results in variables that have different means and standard deviations, but equal ranges.

3.7 Multivariate Acoustic Backscatter Analysis

Multivariate statistics goes a step further by revealing information about the statistical relationship between the two frequencies. Observing and measuring the relationship between acoustic frequencies was crucial for this project as it was necessary to know how backscatter from the seabed varies as a function of frequency. The balanced dual-frequency backscatter datasets were used in these analyses.

To quantify the strength of the relationship between backscatter values from the two frequencies, the Pearson's correlation coefficient (i.e. Pearson's R) was calculated using the balanced backscatter datasets. This statistic provides a measure of the

direction and degree of association between backscatter values from the two frequencies in a manner not influenced by measurement units (Jensen, 2005).

Least squares regression analysis provides a supplement to correlation analysis (McGrew and Monroe, 2000). This technique quantifies the relationship between a set of points by applying a line of best fit, which minimizes the sum of squared distances between data points and the line. Therefore, the overall trend and form of the relationship between two variables are quantitatively summarized. Scatter plots and linear regression lines were displayed for dual-frequency backscatter data in both study areas, sand and gravel seabeds, and by each ISU. The coefficient of determination (r^2) of each linear regression was also calculated. This is a relative index representative of the strength of the relationship between two variables and is the ratio of the explained variation to the total variation (McGrew and Monroe, 2000).

Residual values from the least squares regressions were also calculated and analyzed. Residual values represent the amount of variation in the dependent variable that cannot be explained by the independent variable (McGrew and Monroe, 2000). Hence, residuals are the portion of data that are not explained by the regression model and can therefore provide useful information and indicate where the model does a poor job in predicting values of the dependent variable (i.e. 38 kHz backscatter). The standard error of the estimate (SE) was calculated and this value indicates the relative error, or in other words, the average distance between an observed value and the regression line or model (McGrew and Monroe, 2000). Descriptive statistics and histograms of the residual values from both study areas, sand and gravel seabeds, and each ISU were examined for normal distributions and statistical trends. The residual values were classified using a quantile classification scheme in order to simplify the wide range of values. The spatial distribution of the classified residuals was mapped onto the surficial geology layer.

Additional morphology variables (i.e. depth, slope, and curvature) were plotted as independent variables on a scatter plot, with the residual values as the dependent variable. The residual values detrends the relationship between the two frequencies and therefore represents the variance in the dependent variable (38 kHz) after we account for the variance that is explained by the independent variable (120 kHz). It is anticipated that the additional morphology variables could improve upon the explained variance (i.e. coefficient of variation) of the dependent variable.

3.8 Multivariate Classification

Both unsupervised and supervised classification techniques were used. Unsupervised techniques classify data into groups with no prior knowledge of group membership, and therefore strongly depend on statistical clusters in the data with similar attribute values. Supervised classification techniques delineate classes based on a priori knowledge of group membership. Such techniques commonly take a sub-sample of the data with known group membership and use this information to train the algorithm to classify the data. Supervised classification algorithms can also use clustering techniques.

It would be a challenge to classify all of the ISUs with a high accuracy if near-nadir backscatter was the only input into the classification. Therefore, additional morphology layers were needed in order to measure these morphological features and improve classification accuracy. SBES bathymetry was used to calculate depth and slope values along track. A MBES dataset (8 m pixel size) for each study area were processed in ESRI's ArcGIS software to produce a curvature layer. Curvature is the second order derivative of bathymetry (slope being the first order derivative). The output curvature values were calculated based on a 3x3 (24 m x 24 m) pixel window. The 8 m

pixels can represent the roughness and morphology of certain ISUs (i.e. GH, GRS, GRL) as the scale at which ripples and other bedforms occur range from 2 m to >5 m. Larger magnitude values indicate larger slopes and positive values indicate the surface is convex while negative values indicate the surface is concave (ESRI, 2006). These morphology variables should provide detection of the shape, roughness, and texture of the ISUs in both study areas. Much of the roughness and texture of the seabed (i.e. ripples) are a result of storm energy reaching the seabed, and therefore depth dependent. Slope detects the rate of change in depth and should be most effective for identifying ripples. Curvature will provide measurements of depressions and crests, which again should provide additional information for the detection of ripples, troughs, as well as boulders. Another commonly used morphology variable used in ASC, namely rugosity, was left out here as it was highly correlated with slope. This would therefore break the assumptions of the supervised classification algorithm. The input layers for the multivariate classification were 38 kHz and 120 kHz backscatter, depth, slope, and curvature.

3.8.1 Unsupervised *k*-means Clustering

K-means clustering is a common and effective method for classifying seabed backscatter values in the field of ASC (e.g. Legendre et al., 2002, Courtney et al., 2005). Unsupervised clustering algorithms group a large number of observations or samples into clusters of similar attribute values based on statistically determined criteria. The user must attempt to assign the groups or classes from the output to information classes (in this particular case these information classes are seabed types). Balanced backscatter data from both frequencies and the morphology layers (i.e. depth, slope, and curvature) were combined in a k-means clustering algorithm offered by SPSS in order to classify

the data into similar groups. The scaling of variables is important to consider as the data are compared in multivariate space. If the variables are measured on different scales (e.g. degrees slope, depth in meters, and backscatter in decibels), then they should be standardized (SPSS, 2007). Therefore, the input variables were normalized to a scale from 0 to 1.

K-means requires the user to specify the number of clusters before the algorithm begins, however the optimal number of clusters is not known in this case. The Calinski-Harabasz pseudo F-statistic (C-H) can be used to help identify the optimal number of clusters for the input variables. The C-H algorithm is used for different number of groups using the following equation (Legendre, 2001):

$$C-H = \frac{R^2}{K-1} \bigg/ \frac{[1 - R^2]/(n - K)}{K}$$

where,

$$R^2 = (SST - SSE)/SST$$

SST = total sum of squared distances

SSE = sum of squared distances of the objects to their group's own centroids

K = number of groups

The number of groups that have the highest C-H criterion corresponds to the most compact set of groups, or optimal number of groups (Legendre, 2001).

SPSS requires users to specify the number of clusters that should be identified, as well as the maximum number of iterations the algorithm will run, in this case 30. Based on C-H statistic, the optimal number of clusters, or classes, in the data was three in the preferred study area and two in the non-preferred study area. However, as the goal was to see how well the variables (backscatter, depth, slope, curvature) could classify the ISUs of the interpreted surficial geology map, the number of ISUs in each study area was chosen as the number of clusters for k-means analysis. By setting the

number of clusters based on the number of ISUs, instead of statistically significant clusters, k-means was used as an exploratory statistical technique as opposed to an inferential statistical technique.

The Sand ISU in the preferred study area only had four samples, which is an insignificant percentage of the data and it seemed to cause many misclassifications in the supervised classification (see section 4.5.1). Therefore, the number of clusters in this study area was set to six, instead of seven, to remove the Sand ISU. In the non-preferred study area, two clustering procedures were run due to the bimodal depth distribution, one classification for the dataset with depth < 52m and one for depth > 52m. The observation taken with depth < 52m had five different ISUs, however two of the ISUs had only two or three observations which again is too small a number to attempt to classify. Therefore, the k-means clustering procedure for the depth < 52m dataset was set to three clusters, representing the three ISUs (GRI, Sand, and SM) with a significant number of samples. The same scenario occurs for the dataset with depth > 52m, where the number of clusters was set to five, representing the ISUs with a considerable number of samples (GRS, GRI, GL, GH, SB). There is no overlap in the ISUs between the two datasets (i.e. depth > 52m and depth < 52m).

3.8.2 Supervised Discriminant Function Analysis

Discriminant function analysis was used to assess how well the backscatter and morphology variables (depth, slope, and curvature) could predict the ISUs and which of these variables were optimal in predicting group membership. This is a supervised classification technique because the classification algorithm is trained and developed based on a priori knowledge of typical values of each group (i.e. interpreted sediment units). Discriminant analysis answers two basic research questions: 1) what are the

dimensions along which groups (i.e. ISUs) differ and 2) what are the classification functions that best predict group membership (Tabachnick and Fidell, 2001). There were three different runs of discriminant function classification in both study areas; one used 38 kHz backscatter and the morphology variables, another used only 120 kHz backscatter and the morphology variables, and the final run used backscatter from both frequencies and the morphology variables. The non-preferred study area was split into two datasets (depth <52 m and depth >52 m) for classification in order to account for the bimodal depth distribution.

The statistical software package, SPSS, was used to perform discriminant function analysis on the data. Discriminant function analysis essentially attempts to find linear combinations (discriminant functions) of a set of independent variables that best separate the groups of cases. Each function conveys information on the amount of influence each input variable had on the separation of the groups. Therefore, the user is provided with an idea of the importance each variable plays for group discrimination. The procedure automatically chooses a first function that separates the groups as much as possible. It then chooses a second function, orthogonal to the first, which is both uncorrelated with the first function and provides as much further separation as possible (SPSS, 2007). The first discriminant function provides the best group separation, and typically, only the first one or two discriminant functions reliably discriminate among groups (Tabachnick and Fidell, 2001).

Discriminant function scores for each observation are predicted from the sum of the series of weighted predictors (i.e. variables). The weighting is achieved by discriminant function coefficients, and there is a set of coefficients for each discriminant function. The coefficients indicate how heavily each variable is weighted in order to maximize discrimination of groups (Leech et al., 2005). Each observation receives a

separate discriminant function value for each discriminant function when the values of each variable are inserted into the equations (Tabachnick and Fidell, 2001). To solve for the discriminant function score for the i^{th} discriminant function (D_i), the following equation is used (Tabachnick and Fidell, 2001):

$$D_i = d_{i1}z_1 + d_{i2}z_2 \dots + d_{in}z_n$$

d_i = value of the discriminant function coefficient

n = number of variables

z = score on each variable

The discriminant function coefficients are found by maximizing differences between groups relative to differences within groups (see Tabachnick and Fidell, 2001, for detailed explanation of the calculation of coefficients). The analysis uses ordinary least squares to estimate the discriminant coefficients that minimize the within group sum of squares (Friel, 2009).

Discriminant function analysis allows for classification of cases into groups by developing a classification equation for each group. To get a sense of how well the classification procedure performs, the classification functions are derived from samples from which group membership is known. Then, the classification function is used to classify cases for which group membership is not known. This way, actual group membership is compared to predicted group membership (Tabachnick and Fidell, 2001). The data for each case goes into each group's classification equation, which then provides a classification score for each group for the case. The case is then assigned to the group that had the highest classification score, and the equation is as follows (Tabachnick and Fidell, 2001):

$$C_j = c_0 + c_1X_1 + c_2X_2 + \dots + c_nX_n$$

C_j = a score on the classification function for group j

X = raw score of each variable

c_i = classification function coefficient

c_0 = constant

Essential to the classification procedure are the classification coefficients. These are found using the means of the variables and the pooled within-group variance-covariance matrix (Tabachnick and Fidell, 2001).

4. Results

This chapter presents the results from the analyses. Results of the statistical analyses of backscatter from each frequency independently (i.e. univariate) are highlighted in section 4.1. This is followed by the results of the classification of each frequency independently (section 4.2). Section 4.3 presents the results of analyzing both frequency differences at several spatial scales based on continuous and classified backscatter values. Section 4.4 outlines results of the multivariate analysis of dual-frequency seabed backscatter. Section 4.5 covers the results of the multivariate classifications. Finally, section 4.6 provides a comprehensive summary of the main results.

4.1 Univariate Acoustic Backscatter Analysis

The results presented in this section answers research question 1, concerning the univariate (i.e. single frequency) statistical characteristics of near-nadir acoustic backscatter from each frequency. The input data for the univariate backscatter analysis were the original backscatter datasets, which were not processed to spatially match each 38 kHz and 120 kHz sample. As mentioned in the methods chapter, these backscatter values are based on the logarithm of backscatter values and statistical results formally have no units. Only survey lines 2, 3, 4, and 5 coincided with the ISUs from the interpreted surficial geology map (see Figure 13), and were therefore the only survey lines that could be related to seabed classes. Table 5 displays the number of observations from the acoustic surveys over each ISU in both study areas. The different number of soundings between frequencies for the same ISU was due to the fact that the

38 kHz had a number of erroneous points that were removed in the processing stage or because of the slight offset between soundings of the two frequencies (see section 3.3).

Table 5: Count of acoustic samples over each ISU from both acoustic frequencies. Each sample represents five acoustic pings.

ISU	Non-preferred Study Area				Preferred Study Area			
	38 kHz		120 kHz		38 kHz		120 kHz	
	Count	Percent	Count	Percent	Count	Percent	Count	Percent
GRB	91	3.6	66	3.8	276	16.2	275	16.0
GRS	15	0.9	14	0.8	283	16.7	297	16.4
GRF	77	4.6	90	4.6	96	5.7	90	5.2
GL	123	7.3	140	8.0	31	2.0	36	2.4
GRH	209	12.4	226	13.1	-	-	-	-
Sand	545	32.2	547	31.4	9	0.3	9	0.3
SM	624	36.9	627	36.0	736	43.6	737	43.2
SG	-	-	-	-	99	6.5	99	6.5
SB	37	2.2	40	2.3	-	-	-	-
Total	1691	100.0	1743	100.0	1915	100.0	1529	100.0

Figure 17 and Figure 18 display histograms of backscatter values and reveal unimodal distributions. A multimodal distribution would have been one acoustic indicator of obvious distinct seabed classes. These histograms were considered to have a normal distribution because the skewness values were inside the z values (based on the skewness significance test outlined in the methods) making them non-significant (Table 6). According to Tabachnick and Fidell (2001), "Conventional but conservative (0.01) alpha levels are used to evaluate the significance ...". The skewness value for the 120 kHz frequency was greater than the 38 kHz in the non-preferred study area, and these values were higher than either frequency in the preferred study area.

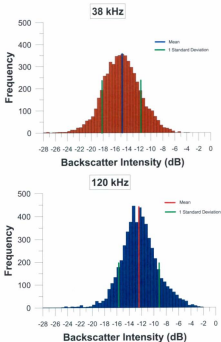


Figure 17: Histograms of near-nadir backscatter values from the non-preferred study area.

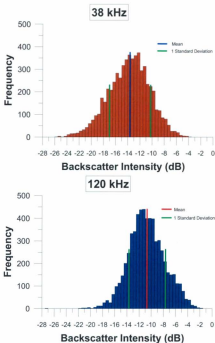


Figure 18: Histograms of near-nadir backscatter values from the preferred study area.

Table 6: Skewness significance test results (S_s = standard error of skewness, z = z distribution score).

Study Area and Frequency	N	Skewness Value	S_s	z	Significant
Preferred 38 kHz	5966	-0.223	0.03171	-7.03248	No
Preferred 120 kHz	6108	0.045	0.03134	1.43386	No
Non-preferred 38 kHz	5398	-0.483	0.03394	-13.88722	No
Non-preferred 120 kHz	5533	-0.891	0.03293	-27.0674	No

Table 7 summarizes descriptive statistics for each frequency in both study areas. The 120 kHz frequency had higher mean near-nadir backscatter values (-12.42 and -10.75) than the 38 kHz frequency (-14.96 and -13.79). Both frequencies in the preferred study area had higher mean backscatter values (-10.75 for the 120 kHz and -13.79 for the 38 kHz) than its corresponding frequency in the non-preferred study area (-12.42 for the 120 kHz and -14.96 for the 38 kHz). A T-test for independent samples was used to compare the mean backscatter values between study areas from the same frequency and results revealed that the means were significantly different ($\alpha = 0.000$ for the 38 kHz and $\alpha = 0.000$ for the 120 kHz) based on a 95% confidence interval. The Levene's test for equality of variances indicated that the variances of the 120 kHz frequency were equal ($\alpha = 0.913$) and the variances for the 38 kHz were not equal ($\alpha = 0.000$) based on a 95% confidence interval. Therefore, a T-test for unequal variances was used for the 38 kHz frequency.

Based on Table 7, the range of values were greater in the non-preferred study area and greater for the 120 kHz frequency. The coefficient of variation (CV) values were also greater for the 120 kHz backscatter values. The 120 kHz frequency in the non-preferred study area had a high kurtosis value (6.28), therefore indicating clustering, while in the preferred study area the 120 kHz frequency had a low kurtosis value (0.14) indicating that the data were more dispersed. The same pattern was true for the 38 kHz frequency. The difference between kurtosis values from the two frequencies was large in

the non-preferred area (3.84) and small in the preferred study area (0.11), with the 120 kHz frequency having the largest values.

Table 7: Descriptive statistics of near-nadir backscatter values from both frequencies in both study areas.

	Non-preferred Study Area		Preferred Study Area	
	38 kHz	120 kHz	38 kHz	120 kHz
N	5308	5533	5966	6108
Minimum	-36.83 dB	-37.71 dB	-25.70 dB	-27.24 dB
Maximum	-3.75 dB	-2.00 dB	-3.80 dB	-1.92 dB
Range	33.08 dB	35.71 dB	21.90 dB	25.32 dB
Mean	-14.96	-12.42	-13.79	-10.75
Standard Deviation	3.25	3.22	3.37	2.96
Variance	10.58	10.38	11.35	8.75
Coefficient of Variation	21.78	25.94	24.43	27.52
Kurtosis	2.44	6.28	-0.03	0.14

Statistics were also calculated for near-nadir backscatter values of sand seabeds (i.e. Sand, SM, SG, and SB) and gravel seabeds (i.e. GRS, GRL, GRI, GL, GH) and presented in Table 8. Independent samples T-test results revealed that mean backscatter values were significantly higher for gravel seabeds than sand seabeds for both 38 kHz ($\alpha = 0.000$) and 120 kHz ($\alpha = 0.000$) frequencies in the non-preferred study area, based on a 95% confidence interval. In the preferred study area, T-test results revealed that mean backscatter values were significantly higher for gravel seabeds than sand seabeds for the 120 kHz ($\alpha = 0.000$), but not significantly higher for the 38 kHz ($\alpha = 0.921$). An unequal variance T-test was used for these analyses, as the variance were assumed not equal based on the Levene's test for equality of variances results. Gravel seabeds also had higher range and variance values than sand seabeds for each frequency in both study areas.

The differences between mean 120 kHz and 38 kHz near-nadir backscatter values were higher for gravel seabeds (3.1 – 4.3) than sand seabeds (1.9 – 2.67). The

magnitude of differences between mean backscatter values of sand and gravel seabeds was much greater for the 120 kHz frequency (1.65 – 1.89) than for the 38 kHz (0.02 – 0.69). Variance values were much higher for gravel seabeds. Coefficient of variation values were also higher for gravel seabeds, and were higher for the 120 kHz frequency compared to the 38 kHz frequency. The skewness values were higher for gravel seabeds for each frequency in both study areas. Finally, kurtosis values were higher for sand seabeds in the non-preferred study area, and just the opposite occurred in the preferred study area. In sum, these results conveyed that gravel seabeds provided significantly higher backscatter values and greater frequency differences in mean backscatter values, 120 kHz backscatter had the largest differences between sand and gravel seabeds, and gravel seabeds produced more variable backscatter values.

Table 8: Descriptive statistics of near-nadir backscatter summarized by dominant sand and gravel seabed substrate types.

	Non-preferred Study Area				Preferred Study Area			
	38 kHz		120 kHz		38 kHz		120 kHz	
	Gravel	Sand	Gravel	Sand	Gravel	Sand	Gravel	Sand
N	485	1206	529	1214	676	839	689	841
Minimum	-27.41 dB	-25.51 dB	-24.09 dB	-22.36 dB	-25.70 dB	-24.40 dB	-21.28 dB	-21.27 dB
Maximum	-3.75 dB	-7.09 dB	-2.27 dB	-5.83 dB	-6.45 dB	-5.44 dB	-2.80 dB	-2.81 dB
Range	23.67 dB	18.41 dB	21.82 dB	16.52 dB	19.25 dB	18.97 dB	18.68 dB	18.48 dB
Mean	-18.11	-14.80	-11.31	-12.96	-14.67	-14.80	-10.37	-12.02
SD	3.83	2.49	3.99	2.07	3.85	3.37	3.32	3.08
Variance	14.64	6.18	15.98	4.29	14.81	11.34	11.01	9.52
CV	27.14	16.82	36.28	16.35	26.24	22.94	32.62	26.71
Skewness	-0.30	-0.88	-0.55	-0.11	-0.36	-0.19	-0.28	0.12
Kurtosis	-0.15	0.40	0.11	0.39	-0.37	-0.15	-0.43	0.17

Table 9 presents the results from the descriptive statistics of near-nadir backscatter values of each ISU. The expected trend of sand seabeds having lower mean backscatter values than gravel seabeds generally held true, however there were exceptions. In the non-preferred study area, the ranking of mean backscatter values of each ISU from lowest to highest for the 120 kHz frequency was as follows: GRI, GRS,

Sand, SB, SM, GH, GRL, and GL. The same ranking for the 38 kHz was GRS, Sand, GRI, SB, GH, SM, GRL, and GL. Two ISUs with dominant gravel substrate, namely GRI and GRS, unexpectedly had some of the lowest mean values for each frequency. ISUs with the highest variance from the 120 kHz frequency were SB, GH and GRS, respectively. The highest variance from the 38 kHz also belonged to SB, followed by GRS and GH.

In the preferred study area, the ranking of mean near-nadir backscatter values of each ISU from lowest to highest for the 120 kHz frequency was as follows: GRI, SM, GRS, SG, GRL, Sand, and GL. The same ranking for the 38 kHz was GRS, SM, GRI, SG, GRL, Sand, and GL. Again, GRI and GRS had some of the lowest mean backscatter values for each frequency. SG, SM, and GRL had the highest variance values from the 120 kHz frequency. GRS, Sand, and SM had the highest variance value from the 38 kHz frequency and SG had the lowest variance value. If the order of mean backscatter values of the ISUs was identical for each frequency within each study area then it would be safe to conclude there was no frequency difference or dependency as a function of ISU. However, these results reveal that the order of ISUs is not the same for each frequency, therefore the two frequencies are responding differently to the seabed.

Table 9: Descriptive statistics of near-nadir acoustic backscatter of each ISU.

ISU	N	Range	Min	Max	Mean	SD	Variance	CV	Skewness	Kurtosis
Non-preferred Study Area										
120 kHz										
GRS	68	17.87	-25.14	-2.27	-13.62	3.68	13.51	26.44	0.83	0.85
GRL	14	6.97	-11.85	-4.68	-8.90	2.30	5.28	25.84	0.43	-1.10
GRI	80	15.21	-22.79	-7.58	-14.07	2.74	7.53	19.47	-0.81	1.74
GL	160	13.57	-18.69	-3.11	-7.63	2.36	5.67	31.19	-0.85	1.94
GH	229	16.23	-24.09	-4.86	-11.30	3.70	13.69	32.74	-1.36	1.53
Sand	547	11.03	-19.85	-8.63	-13.49	1.80	3.25	13.34	-0.36	0.15
SM	627	11.84	-18.47	-6.64	-12.47	2.21	4.94	16.23	-0.19	0.82
SB	40	16.32	-22.36	-5.83	-12.58	3.75	14.05	29.81	-0.42	-0.14
38 kHz										
GRS	61	16.61	-23.80	-7.18	-17.73	3.71	13.78	20.92	0.78	0.32
GRL	15	10.81	-17.86	-7.05	-15.41	3.34	11.16	29.27	-0.62	-0.48
GRI	77	12.17	-22.56	-9.1	-15.63	2.79	7.78	17.85	-0.55	-0.44
GL	123	15.40	-19.15	-3.75	-11.30	2.76	7.63	24.42	-0.25	0.42
GH	259	20.89	-27.41	-6.54	-14.34	3.68	12.10	24.27	-0.57	0.80
Sand	545	16.41	-25.51	-8.09	-15.79	2.13	4.45	13.49	-0.26	0.75
SM	624	14.46	-21.75	-7.29	-13.91	2.31	5.36	16.61	-0.26	0.27
SB	37	16.66	-23.76	-7.09	-15.39	4.08	16.64	26.51	-0.12	-0.62
Preferred Study Area										
120 kHz										
GRS	275	14.66	-19.36	-4.68	-11.71	2.81	7.88	24.00	-0.15	-0.28
GRL	287	17.67	-26.27	-2.60	-8.84	2.84	8.08	32.13	-0.66	0.48
GRI	80	13.82	-21.26	-7.44	-13.13	2.47	6.10	16.81	-0.38	0.52
GL	36	9.47	-13.93	-3.56	-6.93	2.44	5.96	37.37	-1.09	0.89
Sand	5	3.47	-9.72	-6.25	-7.62	1.33	1.76	17.45	-1.14	1.40
SM	737	16.46	-21.27	-2.81	-12.30	2.85	8.08	23.82	0.04	0.23
SG	99	15.27	-18.32	-3.05	-10.24	3.49	12.20	34.08	-0.05	-0.28
38 kHz										
GRS	276	16.09	-26.70	-6.81	-16.42	3.80	14.46	23.14	-0.14	-0.53
GRL	283	15.62	-22.86	-7.26	-13.48	3.23	10.42	24.18	-0.48	-0.12
GRI	86	14.73	-23.25	-8.51	-14.74	3.31	10.94	22.46	-0.53	-0.39
GL	31	12.46	-18.91	-6.45	-10.94	3.30	10.80	31.31	-1.07	0.53
Sand	5	7.75	-15.51	-7.76	-11.53	3.62	13.17	31.48	0.07	-2.85
SM	736	16.87	-24.40	-5.44	-14.85	3.36	11.26	22.63	-0.23	-0.20
SG	88	15.32	-21.39	-6.07	-13.64	3.21	10.26	23.53	0.11	-0.23

Each ISU had distinct mean near-nadir backscatter values. For instance, if a pair of ISUs (e.g. GRS and GRI) had similar mean backscatter values for one frequency, then the other frequency would have quite distinct mean backscatter values (Table 10 and Table 11). For example, in the non-preferred study area, the difference between mean values of GRI and GRS for the 120 kHz frequency was only 0.15. However, the difference between the same two ISUs for the 38 kHz frequency was 2.1. This result

ensures that distinction between ISUs should be easily achievable based on dual-frequency backscatter values as at least one frequency is expected to have large differences. Nevertheless, certain pairs of ISUs had very little difference (< 1) between mean backscatter values for both frequencies. For example, in the non-preferred study area, the difference between the GRI and the sand ISU was 0.58 for the 120 kHz, and 0.16 for the 38 kHz. As a result, these ISUs may be more difficult to differentiate based on backscatter values.

Table 10: Absolute difference in mean near-nadir backscatter values between ISUs in the preferred study area.

120kHz								
	GRS	GRL	GRI	GL	Sand	SM	SB	
GRS	-							
GRL	2.87	-						
GRI	1.72	4.29	-					
GL	5.16	2.31	6.60	-				
Sand	4.09	1.22	5.51	1.09	-			
SM	0.50	3.46	0.83	5.77	4.68	-		
SB	1.47	1.40	2.89	3.71	2.62	2.06	-	
38 kHz								
	GRS	GRL	GRI	GL	Sand	SM	SB	
GRS	-							
GRL	3.32	-						
GRI	1.88	1.34	-					
GL	5.80	2.80	4.20	-				
Sand	4.89	1.87	3.21	0.98	-			
SM	1.51	1.45	0.11	4.33	3.32	-		
SB	2.76	0.24	1.10	3.13	2.11	1.21	-	

Table 11: Absolute difference in mean near-nadir backscatter values between ISUs in the non-preferred study area.

120kHz								
	GRS	GRL	GRI	GL	GH	Sand	SM	SB
GRS	-							
GRL	5.31	-						
GRI	5.18	5.17	-					
GL	6.20	1.27	6.44	-				
GH	3.82	2.4	3.77	3.67	-			
Sand	0.43	4.99	0.98	5.88	2.19	-		
SM	1.51	3.51	1.66	4.78	1.11	1.88	-	
SB	1.34	3.68	1.49	4.95	1.28	0.81	0.17	-
38 kHz								
	GRS	GRL	GRI	GL	GH	Sand	SM	SB
GRS	-							
GRL	6.32	-						
GRI	2.10	4.22	-					
GL	6.43	0.11	4.33	-				
GH	3.29	2.35	1.26	3.04	-			
Sand	1.84	4.38	0.16	4.48	1.45	-		
SM	3.82	2.50	1.72	2.61	0.43	1.88	-	
SB	2.34	3.98	0.24	4.09	1.05	0.40	1.48	-

The non-parametric Kruskal-Wallis H test showed that there were significant differences between the ISUs with regards to ranked backscatter values from each frequency in both study areas (Table 12). The Mann-Whitney U test compared each ISU with every other ISU and revealed whether there were significant differences in ranked backscatter values. This resulted in 28 comparisons in the non-preferred study area and 21 in the preferred study area (Appendix B). Significant differences were based on a widely used p -value of 0.05, meaning if the p -value of the Mann-Whitney U test for a pair of ISUs was less than 0.05 then they were considered significantly different. Based on this threshold, the majority of ISU pairs were significantly different (29 out of 49). Most of the results with no significant differences (13 out of 20) unexpectedly occurred when comparing two ISUs of different dominant substrate types (i.e. when comparing an ISU with dominant sand substrate to an ISU with dominant gravel substrate, or vice versa).

Table 12: Kruskal-Wallis H tests results (see section 3.4 for explanation of mean rank and Chi-square).

ISU	Non-preferred Study Area				Preferred Study Area			
	38 kHz		120 kHz		38 kHz		120 kHz	
	N	Mean Rank	N	Mean Rank	N	Mean Rank	N	Mean Rank
GRS	61	403.06	66	568.42	278	563.18	279	716.15
GRL	15	1299.87	14	1454.57	263	918.15	267	1097.60
GRI	77	675.45	80	578.68	86	757.41	80	918.29
GL	123	1354.99	140	1855.89	31	1227.00	36	1335.80
SH	208	914.02	228	1708.28	-	-	-	-
Sand	545	627.05	547	648.60	5	1894.80	5	1286.80
SM	624	973.03	627	882.74	736	732.58	737	627.42
SB	27	753.78	40	860.65	-	-	-	-
SG	-	-	-	-	88	867.28	89	898.28
Chi-Square	303.363		403.485		148.369		348.211	
Degrees of Freedom	7		7		6		6	
Significance	0.006		0.006		0.006		0.006	

4.2 Univariate Acoustic Backscatter Classification

The remainder of the analyses in this thesis used the balanced datasets. These datasets have the geographic coordinates of each set of dual-frequency soundings averaged to account for the slight delay between soundings from the two frequencies. The resulting datasets have an equal number of 38 kHz and 120 kHz observations and each observation has a 38 kHz and a 120 kHz backscatter value (see section 3.3). Table 13 displays the number of observations over each ISU from the balanced datasets. The first set of analyses using the balanced datasets was the classification of backscatter values from each frequency.

Table 13: Count of observations over each ISU from the balanced datasets in both study areas.

ISU	Non-preferred Study Area			Preferred Study Area		
	Count	Percent	Cumulative Percent	Count	Percent	Cumulative Percent
GRS	61	3.7	3.7	289	18.2	18.2
GRL	14	0.8	4.5	287	19.5	37.7
GRN	75	4.5	9.0	79	5.4	43.1
GL	132	8.0	17.0	34	2.3	45.4
GN	217	13.1	30.1	-	-	-
Sand	519	31.3	61.4	4	0.3	45.7
SM	605	36.9	97.8	708	48.0	93.7
SG	-	-	-	93	6.3	100.0
SR	38	2.2	100.0	-	-	-
Total	1659	100.0		1478	100.0	

The classification of near-nadir backscatter values of each frequency was first performed with a quantile classification method using ESRI ArcGIS software. Classifying with quantiles attempts to see where samples fit in the data distribution for each acoustic frequency (Figure 19). For instance, some samples can be in the right part of the histogram (right of the median), while others can be in the left part. Comparing quantile classes between samples of the same acoustic frequency can indicate relative differences in backscatter responses. The quantile classes ranged from 1 to 5; therefore,

lower backscatter values were represented by lower numbered classes and higher backscatter values were represented by higher numbered classes. How single frequency classification results relate to seabed substrate types is the focus of research question 2 and is answered with the results presented here.

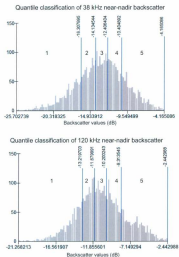


Figure 19: Example of the delineation of the quantile classes (defined by the area between the blue bars) in the preferred study area. Numbers 1 to 5 indicate the quantile classes.

Figure 20 displays the spatial distribution of the 120 kHz frequency quantile classification results overlaid on the ISUs in the preferred study area (refer to Appendix C for the remaining quantile classification maps). Generally, the spatial distribution of the quantile classification results of each frequency revealed that higher numbered quantile classes (i.e. higher backscatter values) related to gravel seabed types, whereas lower numbered quantile classes (i.e. lower backscatter values) related to sand seabed types for both study areas. Outside of this, it was difficult to associate a specific class to certain ISUs based on visual examination of the spatial distribution of quantile classification results.

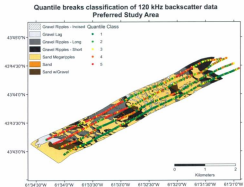


Figure 20: Results of the quantile classification of the 120 kHz near-nadir backscatter values in the preferred study area overlaid on the interpreted sediment units.

The patterns found in the quantile classification results summarized in Table 14 and Table 15 were similar to the those found in the descriptive statistics results of the continuous backscatter values (see section 4.4.1), and need not be repeated here. However, these results indicated that the expected trend of higher backscatter values from gravel seabeds, as opposed to sand seabeds, was more prominent in the 120 kHz frequency than in the 38 kHz frequency. For instance, gravel seabeds (all combined ISUs with dominant gravel substrate) in the preferred study area had more occurrences of low quantile classes than high for the 38 kHz frequency (Table 14). In addition, sand seabeds (all combined ISUs with dominant sand substrate) in the non-preferred study area had more occurrences of high quantile classes than low for the 38 kHz frequency (Table 15). However, only certain ISUs influenced these patterns observed from the 38 kHz frequency. For instance, GRS and GRI in the preferred study area caused the frequent occurrence of low quantile classes. In the non-preferred study area, SM had more occurrences of high quantile classes influencing the result of sand seabeds. The amount of area coverage of the ISUs must also be considered, as SM in the non-preferred study area for example covers much more area than the other ISUs presenting a bias influence when combining all sand seabeds. This leads one to believe that perhaps combining the ISUs into sand and gravel seabeds based on the dominant substrate type of ISUs could be, at times, misleading.

Table 14: Count of observations of quantile classes in the preferred study area.

Quantile Class	Gravel Seabeds	GRS	GRL	GRI	GL	Sand Seabeds	Sand	SM	SO
38 kHz									
1	211	134	55	24	2	235	0	217	18
2	122	49	56	16	2	286	1	135	29
3	134	44	27	20	3	182	0	137	15
4	112	27	67	13	5	133	1	115	16
5	160	15	27	6	22	79	2	62	15
n	689	269	267	79	34	695	4	706	93
120 kHz									
1	135	75	23	36	0	276	0	262	18
2	95	48	23	22	2	181	0	169	12
3	119	58	38	13	1	135	0	117	18
4	138	39	68	7	4	115	0	94	18
5	181	28	138	1	27	88	4	68	26
n	569	268	267	79	34	695	4	706	93

Table 15: Count of observations of quantile classes in the non-preferred study area.

Quantile Class	Sand Seabeds	Gravel Seabeds	GRS	GRL	GRI	GL	GH	Sand	SM	SO
38 kHz										
1	121	85	35	1	17	3	31	87	24	18
2	219	53	11	1	15	4	32	233	61	5
3	272	56	6	0	17	2	31	143	126	3
4	314	91	5	3	18	21	46	119	158	7
5	234	215	8	9	19	102	87	37	186	11
n	1160	499	65	14	75	132	217	519	605	36
120 kHz										
1	226	91	27	0	29	1	34	136	75	8
2	297	46	13	0	22	0	13	167	134	6
3	274	67	6	0	12	2	28	129	141	4
4	356	86	5	6	7	9	42	78	173	5
5	113	244	11	8	5	120	130	19	81	13
n	1160	499	68	14	75	132	217	519	605	36

4.3 Frequency Differences

To assess frequency differences, near-nadir backscatter values from both frequencies were directly compared using several methods. First, the quantile classes from each frequency were differenced in order to assess relative frequency differences based on classified backscatter data. Second, frequency differences of continuous (i.e. unclassified) backscatter values were assessed by directly comparing linear regression lines within certain window sizes (i.e. spatial scales). Third, a statistical difference of means test (Wilcoxon matched-pairs signed-ranks test) was used to test for significant

frequency differences of each ISU based on continuous backscatter values. These results answer research question 3, which asks where the major frequency differences occur and what are the spatial scales at which these frequency differences occur?

4.3.1 Frequency Differences Based on Classified Backscatter Values

The results of the quantile classification of each frequency were differenced in order to explore relative frequency differences based on classified backscatter values. For each sample, the 38 kHz quantile class was subtracted from 120 kHz quantile class, resulting in values ranging from -4 to +4. Consequently, positive values indicated a higher quantile class (i.e. higher relative backscatter) for the 120 kHz backscatter, and negative values indicated a higher quantile class for the 38 kHz backscatter. The larger the result from differencing two quantile classes, the greater the magnitude of difference between the quantile classes from the two frequencies. Zero means no difference in terms of the classified data of the two frequencies.

The results of the differencing in both study areas were normally distributed (Figure 21), meaning smaller magnitudes of difference between the quantile classes were more frequent than larger magnitudes of difference. A normal distribution suggests that neither frequency was dominant in terms of having more occurrences of higher quantile classes.

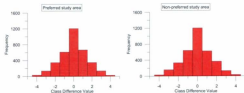


Figure 21: Distribution of the results of differencing the quantile classes (120 kHz - 38 kHz).

In the preferred study area (Table 16), the percentage of samples showing no frequency differences was only 34.3%. 32.3% of the data had a higher quantile class for the 38 kHz frequency, while 33.4% had a higher quantile class for the 120 kHz frequency. The non-preferred area had similar numbers, with 34.1% of the data showing no frequency differences, 32.9% having a higher 38 kHz class, and 33% having a higher 120 kHz class.

Table 16: Count of observations of differencing the quantile classes of each frequency in both study areas (Class Difference = quantile class difference value).

Class Difference	Preferred Study Area			Non-preferred Study Area		
	Count	Percent	Cumulative Percent	Count	Percent	Cumulative Percent
-4	37	1.1	1.1	32	.9	.9
-3	138	3.9	5.0	159	4.5	5.4
-2	346	9.9	14.8	330	9.3	14.8
-1	613	17.5	32.3	542	15.2	30.0
0	1285	34.3	66.6	1263	34.1	64.1
1	695	19.3	85.9	527	14.6	78.7
2	347	9.9	95.8	387	10.4	89.1
3	138	3.9	99.7	134	3.8	92.9
4	28	0.7	100.0	36	1.0	100.0
Total	3570	100.0		3638	100.0	

In the preferred study area, there were roughly even amounts of observations with no differences in the quantile classes of the two frequencies with the sand seabeds having 36.6% no difference and the gravel seabeds having 32.9% no difference (Table 17 and Figure 22). In contrast, the non-preferred study area had 49.7% of the gravel

seabeds and 25.7% of the sand seabeds showed no difference in quantile classes. Therefore, there were clearly more differences in quantile classes of the two frequencies over sand seabeds in the non-preferred study area.

Table 17: Count of observations of differencing the quantile classes from each frequency summarized by sand and gravel seabed types (Class Difference = quantile class difference value).

Class Difference	Count	Percent	Cumulative Percent	Count	Percent	Cumulative Percent
	Sand Seabed Types			Gravel Seabed Types		
			Preferred Study Area			
-4	11	1.4	1.4	4	0.6	0.6
-3	29	3.6	5.0	11	1.6	2.2
-2	87	8.3	13.3	26	3.9	6.1
-1	154	18.1	31.4	76	11.4	17.5
0	295	36.6	68.1	220	32.9	50.4
1	149	18.5	87.6	160	25.1	75.5
2	70	9.4	97.0	104	15.5	91.0
3	21	2.6	99.6	52	7.8	98.8
4	3	0.4	100	8	1.2	100.0
Total	885	100		689	100	
			Non-preferred Study Area			
-4	18	1.6	1.6	7	1.4	1.4
-3	34	6.1	7.7	17	3.4	4.8
-2	178	15.0	24.7	30	6.0	10.8
-1	302	26.0	50.7	64	12.8	23.6
0	298	25.7	76.4	246	48.7	72.3
1	161	13.9	90.3	72	14.4	87.8
2	77	6.6	96.9	36	7.2	95.0
3	31	2.7	99.6	18	3.6	98.6
4	6	0.4	100.0	7	1.4	100.0
Total	1191	100.0		499	100.0	

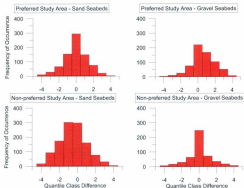


Figure 22: Distribution of the results of differencing the quantile classes (120 kHz - 38 kHz) summarized by sand and gravel seabeds in both study areas.

Figure 23 and Figure 24 map the quantile class differences on the ISUs for both study areas. A red point indicates that the 38 kHz quantile class was higher than the 120 kHz class, a blue point shows that the 120 kHz quantile class was higher than the 38 kHz class, and a black point represents no quantile class differences. These figures indicated that 38 kHz quantile classes were generally higher than 120 kHz quantile classes over sand seabeds. On the other hand, gravel seabeds generally had higher quantile classes for the 120 kHz frequency. Table 17 conveyed that sand seabeds in the preferred study area had more or less a normal distribution of values from the differencing of quantile classes, while the gravel seabed types had more occurrences of

higher 120 kHz quantile classes. Conversely, in the non-preferred study area, gravel seabed types had a normal distribution of values from the differencing of quantile classes, while the sand seabed types had more occurrences of higher 38 kHz quantile classes (Table 17).

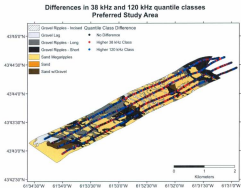


Figure 23: Results of the differences between quantile classifications of each frequency for the preferred study area.

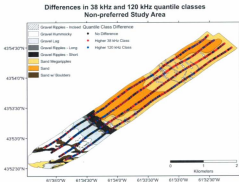


Figure 24: Results of the differences between quantile classifications of each frequency for the non-preferred study area.

Table 18 and Table 19 present the distribution of quantile class differences for each ISU. Certain ISUs did not follow the expected pattern of gravel seabeds having higher quantile classes from the 120 kHz and sand seabeds having quantile classes from the 38 kHz. For both study areas, GRU had more negative values than positive ones. SG from the preferred study area, and SB from the non-preferred study area, had more values that were positive.

Table 18: Count of observations of differencing the quantile classes from each frequency summarized by the ISUs in the non-preferred study area (CD = quantile class difference, F = frequency of occurrence).

CD	GRS		GRL		GRI		GL		GH		Sand		SM		SO	
	F	%	F	%	F	%	F	%	F	%	F	%	F	%	F	%
-4	-	-	-	-	3	4.0	-	-	4	1.8	6	1.2	12	2.0	-	-
-3	1	1.6	-	-	6	10.7	1	0.8	7	3.2	43	8.3	50	8.3	1	2.8
-2	3	4.9	-	-	10	13.3	1	0.8	16	7.4	79	15.2	83	15.4	2	5.6
-1	6	13.1	4	28.6	19	25.3	4	3.0	29	13.4	126	24.3	172	28.4	4	11.1
0	33	54.1	5	35.7	17	22.7	191	26.5	92	42.4	115	22.2	164	27.1	19	52.8
1	6	13.1	3	21.4	14	18.7	17	12.9	30	13.8	89	17.1	86	15.0	3	16.7
2	4	6.6	1	7.1	1	1.3	2	1.5	28	12.9	38	7.3	37	6.1	2	5.6
3	3	4.9	1	7.1	3	4.0	4	3.0	7	3.2	39	7.7	32	5.7	2	5.6
4	1	1.6	-	-	-	-	2	1.5	4	1.8	4	0.8	1	0.2	-	-
Total	61	100	14	100	75	100	132	100	217	100	519	100	665	100	34	100

Table 19: Count of observations of differencing the quantile classes from each frequency summarized by the ISUs in the preferred study area (CD = quantile class difference, F = frequency of occurrence).

CD	GRS		GRL		GRI		GL		Sand		SM		SO	
	F	%	F	%	F	%	F	%	F	%	F	%	F	%
-4	1	0.4	1	0.3	2	2.5	-	-	-	-	11	1.6	-	-
-3	5	1.9	-	-	6	7.6	-	-	-	-	29	4.1	-	-
-2	9	3.0	9	3.1	9	11.4	-	-	-	-	62	8.8	5	5.4
-1	27	10.0	29	10.1	18	22.8	2	5.9	-	-	141	19.9	13	14.0
0	99	35.1	63	27.9	29	36.7	22	64.7	2	60.0	298	36.8	34	36.0
1	69	25.7	79	27.5	14	17.7	6	17.6	1	25.0	124	17.5	24	25.8
2	46	16.7	55	19.2	1	1.3	3	8.8	-	-	64	9.0	12	12.9
3	22	8.2	29	10.1	-	-	1	2.9	1	25.0	17	2.4	3	3.2
4	3	1.1	5	1.7	-	-	-	-	-	-	1	0.1	2	2.2
Total	289	100	287	100	79	100	34	100	4	100	758	100	93	100

Another method used to assess the differences between frequencies based on classified backscatter was to consider the spatial continuity of adjacent points with similar quantile class difference values (no difference, higher quantile class for the 38 kHz, and higher quantile class for the 120 kHz). Results in Table 20 and Table 21 revealed that the most frequently occurring number of adjacent points with the same value was one, and the spatial scale of this translates into the footprint size of one observation, which ranges from 14 m to 21 m in along-track distance depending on depth (see section 3.3 for footprint size calculation). Figure 25 provides an example of a histogram displaying the spatial continuity of points with no frequency differences in the non-preferred study area based on classified backscatter. There was a general decrease in the number of observations with no frequency differences, until approximately eight

continuous observations; at this point, there were some slight increases. There was a general pattern in the remainder of the histograms (Appendix E), where there tended to be a slight increase in the number of observations of continuous points with the same value anywhere between six to twelve observations. This roughly translates into along-track distances ranging from 74m to 153m (12 m is the average spacing between observations and the footprint length along track ranges from 14 m to 21 m depending on depth).

In the preferred study area (Table 20), the maximum along-track distance of the continuity of adjacent observations with no frequency differences based on classified backscatter values approximately ranged from 242 m – 249 m, 122 m – 129 m for higher 38 kHz quantile classes, and 146 m – 153 m for higher 120 kHz quantile classes. Examining the maximum distance of continuity over sand seabeds, the along-track distances ranged from 122 m – 129 m for no frequency differences, 86 m – 93 m for higher 38 kHz quantile classes, and 98 m – 105 m for higher 120 kHz quantile classes. For gravel seabeds the maximum distance of continuity for no differences was 110 m – 117 m, 74 m – 81 m for higher 38 kHz quantile classes, and 122 m – 129 m for higher 120 kHz quantile classes. In comparison, in the non-preferred study area (Table 21), maximum continuity for no frequency differences was 146 m – 153 m, 122 m – 129 m for higher 38 kHz quantile classes, and 182 m – 189 m for higher 120 kHz quantile classes. For sand seabeds in the non-preferred study area, the maximum distance of continuity for no differences was 74 m – 81 m, 122 m – 129 m for higher 38 kHz quantile classes, and 110 m – 117 m for higher 120 kHz quantile classes. For gravel seabeds in the non-preferred study area, the maximum distance of continuity for no differences was 146 m – 153 m, 50 m – 57 m for higher 38 kHz quantile classes, and 62 m – 69 m for higher 120 kHz quantile classes.

Table 20: Summary of the spatial continuity of points from quantile class differencing in the preferred study area. Summarized by no difference, higher quantile class for the 38 kHz frequency (+ 38), and higher quantile class for the 120 kHz frequency (+ 120).

Number of adjacent points with same value	Whole Study Area			Sand Seabeds			Gravel Seabeds		
	No diff	+ 38 kHz	+ 120 kHz	No Diff	+ 38 kHz	+ 120 kHz	No Diff	+ 38 kHz	+ 120 kHz
1	440	260	432	130	102	136	88	54	69
2	136	131	141	40	19	33	17	18	34
3	47	60	47	15	19	10	9	4	16
4	21	41	17	4	6	4	5	4	5
5	12	6	12	8	1	3	0	0	6
6	5	6	9	1	2	2	2	1	6
7	4	4	4	1	2	1	0	0	3
8	6	2	4	1	0	1	1	3	3
9	3	2	3	0	0	3	1	0	1
10	2	4	1	1	0	0	0	3	1
11	1	0	0	0	0	0	0	0	0
12	0	0	1	0	0	0	0	0	0
13	0	0	1	0	0	0	0	0	0
14	0	0	1	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0
18	1	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0
20	1	0	0	0	0	0	0	0	0
n	679	616	673	171	161	167	123	82	144

Table 21: Summary of the spatial continuity of points from quantile class differencing in the non-preferred study area. Summarized by no difference, higher quantile class for the 38 kHz frequency (+ 38), and higher quantile class for the 120 kHz frequency (+ 120).

Number of adjacent points with same value	Whole Study Area			Sand Seabeds			Gravel Seabeds		
	No diff	+ 38 kHz	+ 120 kHz	No Diff	+ 38 kHz	+ 120 kHz	No Diff	+ 38 kHz	+ 120 kHz
1	442	424	480	167	145	168	46	51	72
2	146	154	149	42	12	32	26	21	14
3	41	58	54	9	36	13	11	3	8
4	33	28	17	4	12	1	6	1	1
5	20	9	11	8	8	0	4	0	1
6	5	7	5	1	6	0	1	0	0
7	2	4	3	0	2	0	1	0	0
8	7	1	2	8	1	0	2	0	0
9	0	2	1	8	2	1	0	0	0
10	3	1	1	0	1	0	3	0	0
11	0	0	0	0	0	0	0	0	0
12	1	0	0	0	0	0	1	0	0
13	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0
15	0	0	1	0	0	0	0	0	0
n	666	664	724	223	283	218	101	76	96

Histogram of spatial continuity of samples with no difference between quantile classes

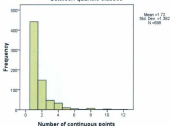


Figure 25: Histogram displaying the frequency of spatial continuity of points with no differences between quantile classes in the non-preferred study area.

4.3.2 Frequency Differences Based on Continuous Backscatter Values

Line graphs (Figure 26 and Figure 27) illustrate examples of near-nadir backscatter intensity from both 38 kHz and 120 kHz frequencies along survey line 2 from both study areas (see Appendix E for remaining survey line graphs). The data were smoothed using a running average of nine observations. The spatial scale of nine observations roughly ranges from 110 m – 116 m along track, as the average spacing between points is 12 m and the footprint length along track ranges from 14 m to 21 m depending on depth. Averaging over nine observations effectively reduced small-scale variability and noise while at the same time maintained the major trends. Values less than nine did not provide enough reduction in small-scale variability.

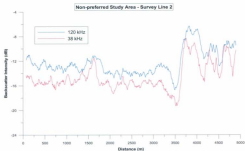


Figure 26: Line graph of near-nadir backscatter values from both frequencies in the non-preferred study area along survey line 2 (surveyed from northeast to southwest). Data were smoothed using a running average of 9.



Figure 27: Line graph of near-nadir backscatter values from both frequencies in the preferred study area along survey line 2 (surveyed from northeast to southwest). Data were smoothed using a running average of 9.

Spatial variations were observed at several spatial scales. By examining the near-nadir backscatter values across the whole survey line (5 km) it was obvious that the 120 kHz frequency backscatter values were higher than the 38 kHz frequency, a result observed with the descriptive statistics (see section 4.1). It was also apparent, by observing trends of backscatter values at this scale of kilometres, that both frequencies followed the same general pattern.

At a scale of hundreds to thousands of meters, two acoustic classes were evident characterized by high and low near-nadir backscatter values. High and low backscatter values could generally be related to sand (lower backscatter response) and gravel (higher backscatter response) seabed types in the two study areas. The cut-off points for discriminating between high and low backscatter values were approximately -11 dB for the 120 kHz and -14 dB for the 38 kHz in the non-preferred study area, and -10 dB for the 120 kHz and -14 dB for the 38 kHz in the preferred study area. These values were based on subjective visual examinations of the line graphs as well as the mean values for seabed types with dominant sand and gravel substrate types as presented in section 4.1. Based on visual interpretations of the line graphs in Figure 26 and Figure 27, the approximate spatial scales of high (700 – 1200 m, 2700 – 3800 m, 4300 – 5000 m) and low (0 – 700 m, 1200 – 2700 m, 3800 – 4300 m) backscatter in the preferred area ranged from 500 m to 1000 m for the 120 kHz frequency. In comparison, the non-preferred area displayed approximate spatial scales ranging from 1000–3000 m for high (3500 – 5000 m) and low (0 – 3500 m) backscatter for the 120 kHz frequency. Therefore, the spatial scales at which these classes occurred seemed to vary between the two study areas. The boundaries between high and low backscatter values were less obvious and more difficult to define for the 38 kHz frequency. This relates to the results in section 4.1 where the differences between mean backscatter values of two dominant

substrate types of the ISUs (sand and gravel) were less obvious for the 38 kHz frequency, especially in the preferred study area. Table 22 reflects this, as there are more observations of high and low backscatter for the 38 kHz frequency meaning there was more variability between high and low backscatter areas. The mean and range distance values from this table also further supports the fact that high and low backscatter areas occur at scales of hundreds to thousand of meters.

Table 22: Descriptive statistics of the along-track distances (meters) of high and low near-nadir backscatter areas in both study areas. High and low backscatter values were defined by the following cut-off values: -11 dB (120 kHz) and -14 dB (38 kHz) in the non-preferred study area and -10 dB (120 kHz) and -14 dB (38 kHz) in the preferred study area (BS = backscatter).

	N	Min	Max	Range	Mean	Standard Deviation	Variance
Non-preferred Study Area							
High BS 38 kHz	41	70.39	1304.15	1233.76	173.80	289.70	83928.56
Low BS 38 kHz	41	10.21	1960.95	1950.74	302.34	421.11	177346.55
High BS 120 kHz	29	9.85	674.76	665.71	134.83	195.66	38274.79
Low BS 120 kHz	31	0.54	3575.39	3565.85	503.96	944.64	892336.31
Preferred Study Area							
High BS 38 kHz	50	8.29	802.80	794.51	167.64	193.56	37475.42
Low BS 38 kHz	50	8.30	2631.73	2623.43	223.71	479.36	175864.64
High BS 120 kHz	23	10.27	1149.96	1139.68	237.33	325.96	106244.03
Low BS 120 kHz	24	11.29	4812.83	4801.54	587.87	987.17	1074897.217

Examining the patterns more closely, at scales of tens to hundreds of meters, it is noticeable that there are areas along the survey line where near-nadir backscatter intensity shows diverging signals between frequencies for the same area of the seabed. Hence, examining backscatter at a more local scale reveals that there are divergences in the backscatter response of specific sampled areas as a function of frequency. Therefore, there are geological properties in the seabed that responded differently to the two frequencies.

To extend on examining the local scale variability of these smoothed near-nadir backscatter data, areas of frequency divergence were objectively identified by directly comparing continuous backscatter values from each frequency. The slope values of the linear regression lines from both frequencies were directly compared within different

moving windows as explained in section 3.6. The spatial scales along track of each of these window sizes varied. Because each smoothed data sample represented along-track distances of 110 m – 116 m and the average distance between points is 12 m, then the spatial scale of a 5-sample window represented a distance of 158 m – 164 m. Furthermore, a 15-sample window ranged from 278 m – 284 m, 25 ranged from 398 m – 404 m, and 35 ranged from 518 m – 524 m. This approach allowed for the examination of divergences at different scales.

Table 23 shows the percentage of samples with divergences for each study area, as well as for the sand and gravel seabeds within each study area, keeping in mind that the amount of gravel and sand seabeds are different within each study area. The percentages of samples with divergences in each study area decreased with increasing window size. An independent samples T-test indicated that the percentage of divergences for sand seabeds were statistically similar between study areas ($\alpha = 0.193$), but differed for gravel seabeds ($\alpha = 0.025$), based on a 95% confidence interval and $N = 4$. The trend of less frequency divergences in the non-preferred study area for gravel seabeds suggests that there was less variation in the backscatter response of gravel seabeds as a function of frequency. These divergence areas were mapped over the interpreted surficial geology maps (Figure 28 and Appendix E).

Table 23: Percentages of samples with frequency divergences based on certain window sizes.

Window Size	Non-preferred Study Area			Preferred Study Area		
	Whole Study Area	Sand Seabeds	Gravel Seabeds	Whole Study Area	Sand Seabeds	Gravel Seabeds
5	35.3	38.4	25.7	32.8	31.8	33.2
15	29.8	32.6	23.3	30.3	31.8	31.2
25	27.2	32.0	18.8	27.8	36.8	29.0
35	22.8	30.2	13.8	24.4	28.8	25.1

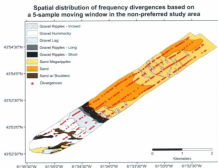


Figure 28: Areas of frequency divergence within a 5-sample moving window overlaid onto the interpreted surficial geology map in the non-preferred study area.

The Wilcoxon matched-pairs signed-ranks test indicated whether there were significant differences between backscatter values from the two acoustic frequencies for each ISU normalized to a 0 to 1 scale. Significant frequency differences were based on a 0.05 p-value threshold. For the non-preferred study area (Table 24), there were significant frequency differences for the following ISUs: GRS, GRI, GL, and SM. The following ISUs had significant frequency differences in the preferred study area (Table 25): GRS, GRI, GRL, and SM. GL was also very close to the 0.05 p-value. From these results, it is obvious that GRS, GRI, GL, and SM had quite distinct backscatter values between frequencies in both study areas. If one considers the rank of the ISUs according to the significance level, one could conclude that SM and GL were most significantly different in terms of backscatter values from both acoustic frequencies, followed by GRS

and GRI, in the non-preferred study area. Interestingly, there were no differences in the rank of significance values for the ISUs with significant differences in the preferred study area. Only SM had significant frequency differences of all ISUs with dominant sand substrate. Because the Sand ISU did not show significant frequency differences, it can be inferred that small scale structure of the seabed caused frequency differences and not the sediment itself.

Table 24: Non-preferred study area Wilcoxon matched-pairs signed-ranks test of near-nadir backscatter values from each frequency for each ISU (see section 3.6 for explanation of sum of ranks used, mean rank, sum of ranks, and Z).

ISU	Study Area Coverage (%)	Sum of Ranks Used	N	Mean Rank	Sum of Ranks	Z	Significance
GRS	8.0	Negative	21	27.38	575	-2.661	0.008
GRI	3.8	Negative	6	8.80	51	-0.099	0.925
GRD	3.7	Positive	30	32.77	983	-2.334	0.020
GH	16.7	Negative	96	115.73	10630	-1.292	0.196
GL	7.8	Negative	45	58.64	2645	-3.954	0.000
Sand	26.2	Positive	243	254.04	61725	-1.879	0.060
SM	37.5	Positive	212	264.88	56154	-6.254	0.000
SG	2.4	Negative	18	17.81	285	-0.754	0.451

Table 25: Preferred study area Wilcoxon matched-pairs signed-ranks test of near-nadir backscatter values from each frequency for each ISU (see section 3.6 for explanation of sum of ranks used, mean rank, sum of ranks, and Z).

ISU	Study Area Coverage (%)	Sum of Ranks Used	N	Mean Rank	Sum of Ranks	Z	Significance
GRS	17.3	Negative	86	119.03	10476	-8.015	0.000
GRI	14.9	Negative	94	119.63	11265	-8.683	0.000
GRD	4.9	Positive	22	25.41	559	-4.980	0.000
GL	3.2	Negative	11	16.85	188	-1.906	0.057
Sand	3.6	Negative	2	2	4	-0.385	0.715
SM	52.8	Positive	268	212.15	63281	-7.754	0.000
SG	8.7	Negative	44	45.30	1903	-3.738	0.001

4.4 Multivariate Acoustic Backscatter Analysis

One of the objectives of this study was to run a classification using backscatter data from each frequency as the two input variables, thereby resulting in a multivariate classification. The basic assumption is that the two frequencies contain different information about similar areas and therefore may improve classification when combined. The multivariate analysis presented in this section answers research

question 4, which inquires about the statistical characteristics of dual-frequency near-nadir backscatter. Pearson's correlation coefficient (i.e. Pearson's R) quantifies and conveys information about the direction and degree of association near-nadir between backscatter values from both frequencies. The Pearson's R correlation results revealed that there was a significant (<0.01) positive correlation between 38 kHz and 120 kHz backscatter values in both study areas ($R = 0.56$ in the preferred study area and $R = 0.52$ in the non-preferred study area). The correlation values of 0.56 and 0.52 suggest that despite the general agreement between both frequencies, the degree of association is not particularly strong.

Figure 29 and Figure 30 indicate no obvious clusters that could allow the delineation of specific seabed types. A least-squares linear regression line characterized the relationship between the two variables (120 kHz and 38 kHz backscatter values) and revealed information about the overall trend and form of the relationship. The linear regression equation for the non-preferred area is $Y = -7.5764 + 0.5569 * X$, and the preferred area equation is $Y = -7.7124 + 0.6087 * X$. The regression coefficient values indicated that a general increase in near-nadir backscatter values from the independent variable (120 kHz) corresponded to an increase in backscatter values from the dependent variable (38 kHz). The coefficient of determination (r^2) values indicated that only 27% (non-preferred study area) and 31% (preferred study area) of the variation in the 38 kHz backscatter values could be explained by the 120 kHz backscatter values. Therefore, there is a large unexplained variance in the backscatter values from the two frequencies. However, based on F-tests that examined whether the regressions explained a statistically significant portion of the variability in the dependent variable from variability in the independent variable, the significance values of the two linear

regressions were less than 0.05, meaning that the variation explained by the models is still significant (SPSS, 2007).

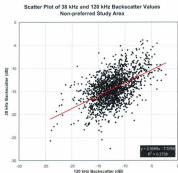


Figure 29: Scatter plot of near-nadir backscatter values for the four survey lines that overlap the interpreted surficial geology map in the non-preferred study area.

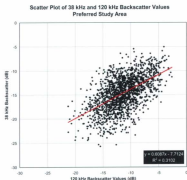


Figure 30: Scatter plot of near-nadir backscatter values for the four survey lines that overlap the interpreted surficial geology map in the preferred study area.

Scatter plots (Figure 31, Figure 32, and Appendix F) of near-nadir acoustic backscatter from both the 38 kHz and 120 kHz frequencies were summarized by sand seabeds (i.e. Sand, SM, SG, SB) and gravel seabeds (i.e. GRS, GRL, GRl, GL, GH). Based on the regression outputs in SPSS, a T-test used to compare the regression coefficients (i.e. slope) of the two dominant sediment types revealed that the coefficients were statistically different based on a 0.05 significance level. In other words, the two dominant sediment types produced varying backscatter as the regression models that were fit to the data had statistically different slopes.

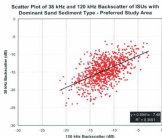


Figure 31: Scatter plot of dual-frequency near-nadir backscatter intensities of ISUs with dominant sand sediment type in the preferred study area. The regression line is the one defined for all the backscatter data that overlapped the ISUs.

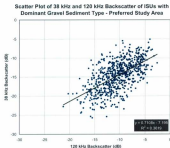


Figure 32: Scatter plot of dual-frequency near-nadir backscatter intensities of ISUs with dominant gravel sediment type in the preferred study area. The regression line is the one defined for all the backscatter data that overlapped the ISUs.

Scatter plots were further processed to display the corresponding ISUs from the interpreted surficial geology classification for each data point (Figure 33 and Appendix F). Figure 33 shows no distinct clusters of ISUs that clearly diverge from the regression line. Divergences from the regression line would have indicated that certain ISUs could be effectively identified using the two frequencies. An effective way of quantitatively examining whether certain ISUs tend to cluster above or below the linear regression line of all the near-nadir backscatter values is to examine the residuals.

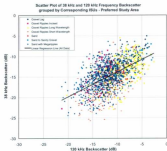


Figure 33: Scatter plot of dual-frequency near-nadir acoustic backscatter intensities symbolized by interpreted sediment units in the preferred study area.

The residuals, or unexplained variances, of the dependent variable (38 kHz backscatter) were calculated from the linear regression model for each study area. The standard error of the estimate (SE) of the residuals was 2.94 for the preferred study area and 2.85 for the non-preferred study area. Histograms of residual values from each study

area are displayed in Figure 34 and Figure 35. These graphs illustrate the normal distribution of the backscatter residual values. Figure 36 and Figure 37 displays scatter plots of the 38 kHz residuals and 120 kHz backscatter values from both study areas. Both scatter plots display homoscedasticity, meaning that there are roughly an equal number of points on either side of the regression line.

Residuals of the linear regression for 38 kHz and 120 kHz backscatter values - Non-preferred Study Area

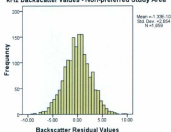


Figure 34: Histogram of the backscatter residual values from the 38 kHz and 120 kHz linear regression model in the non-preferred study area.

Residuals of the linear regression for 38 kHz and 120 kHz backscatter values - Preferred Study Area

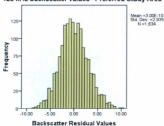


Figure 35: Histogram of the backscatter residual values from the 38 kHz and 120 kHz linear regression model in the preferred study area.

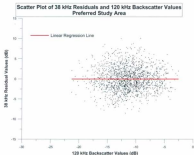


Figure 36: Scatter plot displaying homoscedasticity of 38 kHz residuals and 120 kHz backscatter in the preferred study area.

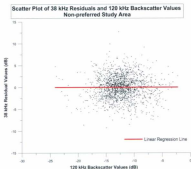


Figure 37: Scatter plot displaying homoscedasticity of 38 kHz residuals and 120 kHz backscatter in the non-preferred study area.

Table 26 summarizes the descriptive statistics of residuals values for sand seabeds, gravel seabeds, and each ISU. In the non-preferred area, the mean residual values for sand (0.02) and gravel (-0.04) were low, whereas in the preferred study area, the values for sand (-0.41) and gravel (0.54) were much higher. In terms of the ISUs, GRS (2.12), GRL (-1.26), and GL (-1.06) had rather large mean residual values in the non-preferred study area. In the preferred study area, GL (-1.64) and GRS (1.55) again had high mean residual values, along with GRL (-0.93). The linear regression model, therefore, does a relatively poor job predicting 38 kHz backscatter values for these ISUs.

Table 26: Descriptive statistics of residual values.

ISU	N	Range	Min	Max	Mean	SD	Variance	CV	Skewness	Kurtosis
Non-preferred Study Area										
Gravel Seabeds	489	21.96	-9.22	12.74	-0.04	3.25	10.21	0.000	-0.028	0.138
Sand Seabeds	1180	17.42	-7.48	9.95	0.02	2.38	5.69	0.1958	0.123	0.180
GRS	61	13.76	-5.57	8.19	2.12	2.05	8.70	139.15	-0.466	-0.137
GRL	14	18.01	-5.58	4.43	-1.26	3.25	10.58	257.84	0.348	-1.126
GRI	75	18.28	-4.61	5.67	0.38	2.60	6.76	886.87	0.034	-0.581
GL	132	13.44	-7.28	8.17	-1.06	2.87	8.22	270.75	-0.028	-0.208
GH	217	21.96	-9.23	12.74	-0.06	3.33	11.09	8590	-0.028	0.888
Sand	519	15.24	-5.30	9.95	0.79	2.27	5.15	324.29	0.182	0.277
SM	605	15.12	-7.48	7.65	-0.60	2.28	5.13	376.87	0.103	0.078
SR	38	13.38	-6.43	8.94	0.33	3.19	10.18	966.87	-0.83	-0.134
Preferred Study Area										
Gravel Seabeds	669	18.08	-8.04	10.04	0.04	3.13	9.81	579.83	0.168	-0.130
Sand Seabeds	805	18.01	-6.55	7.55	-0.41	2.78	7.72	676.85	0.179	-0.020
GRS	269	17.71	-7.67	10.04	1.55	3.18	10.13	295.16	-0.063	-0.132
GRL	287	16.64	-8.04	8.60	0.26	2.67	6.25	1169	0.328	0.125
GRI	79	12.12	-7.65	5.08	-0.93	2.78	7.72	296.82	-0.080	-0.446
GL	34	11.28	-7.36	3.84	-1.84	2.71	7.37	189.24	0.179	-0.131
Sand	4	7.11	-4.16	2.95	-1.26	3.29	10.81	281.11	0.737	-1.669
SM	708	18.10	-8.55	7.55	-0.42	2.84	8.07	676.19	0.173	-0.032
SG	93	10.94	-6.44	5.10	-0.28	2.25	5.04	803.57	0.198	-0.476

Spatial patterns of residuals were analyzed by overlaying residual values on the interpreted surficial geology map. The residuals values were classified into five classes using a quantile classification scheme; much like what was done with the raw backscatter values. This simplifies the large range of the residual values. In contrast to the quantile classification of the raw backscatter values, quantile class values of one and two represent relatively larger negative residual values, whereas quantile class values of four and five are synonymous to relatively larger positive residual values. A quantile

class of three indicates very low residual values, which can include both positive and negative values. Quantile classes are useful for displaying the spatial distribution of residual values over the ISUs of each study area (Figure 38 and Figure 39).

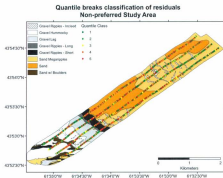


Figure 38: Residual values classified into five quantile classes overlaid onto the interpreted surficial geology map.

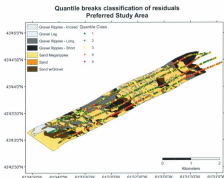


Figure 39: Residual values classified into five quantile classes overlaid onto the interpreted surficial geology map.

These spatial patterns were quantified by examining the number of observations of classified residual values for sand and gravel seabeds, as well as each ISU (Table 27 and Table 28). Large residual values, both positive and negative, occurred more often than small ones over gravel seabeds in the non-preferred study area. The preferred study area had more positive residual values than negative ones over gravel seabeds and more negative residual values than positive ones over sand seabeds. In both study areas, GL and SM had a higher number of negative residual values, whereas GRS and Sand had more positive residual values. GRI had more positive residual values in the non-preferred study area and more negative values in the preferred study area.

Table 27: Count of observations of residual quantile classes for the preferred study area.

Quantile Class	Sand Seabeds	Gravel Seabeds	GRS	GRL	GRI	GL	Sand	SM	SG
1	95	61	17	26	12	6	2	87	6
2	244	144	33	70	26	15	0	215	29
3	235	182	67	84	24	7	1	205	29
4	165	170	85	71	15	5	1	139	25
5	66	106	67	36	2	1	0	62	4
n	805	669	269	287	79	34	4	708	93

Table 28: Count of observations of residual quantile classes for the non-preferred study area.

Quantile Class	Sand Seabeds	Gravel Seabeds	GRS	GRL	GRI	GL	GH	Sand	SM	SB
1	211	121	6	6	14	43	52	52	153	6
2	249	82	5	3	12	27	35	95	145	9
3	247	86	6	0	14	25	41	114	129	4
4	244	88	12	2	17	23	34	126	111	7
5	209	122	32	3	18	14	55	132	87	10
n	1160	499	61	14	75	132	217	519	605	36

The residual values were plotted, as the dependent variable, against morphology layers (independent variable) to determine the relationship between the backscatter residuals and the shape and structure of the seabed. The scatter plots for the backscatter residuals and three morphology variables (depth, slope, and curvature) are displayed in Figure 40 and Figure 41. However, the coefficient of determination (r^2) values were rather small for the linear regression model of the 38 kHz residuals and these morphology variables in both the non-preferred study area (depth (0.0004), slope (0.0001), curvature (0.000003)) and the preferred study area (depth (0.0030), slope (0.0002), curvature (0.0027)). These values suggest that these variables do little to improve upon explained variance of the residuals.

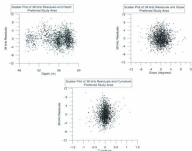


Figure 40: Scatter plots displaying the relationship between morphology layers and backscatter residuals in the preferred study area. r^2 values are small for depth (0.0004), slope (0.0001) and curvature (0.00003).

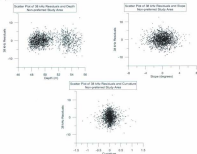


Figure 41: Scatter plots displaying the relationship between morphology layers and backscatter residuals in the non-preferred study area. r^2 values are small for depth (0.003), slope (0.0002) and curvature (0.0027).

4.5 Multivariate Classification

Both supervised (discriminant function) and unsupervised (k-means) methods were used for classifying acoustic data. Both methods used dual-frequency near-nadir backscatter and morphology variables as inputs (see section 3.8 for an explanation behind the selection of morphology variables). Supervised classification attempts to classify the actual seabed types or ISUs, whereas unsupervised clustering provides a number of clusters in multivariate space which are then associated with certain seabed types. The following results addresses research question 5 concerning the relation between seabed substrate types and multivariate classification results, and research question 6 which questions whether dual-frequency (multivariate) classifications improve upon single frequency classification of seabed substrate types.

4.5.1 Supervised Discriminant Function Analysis

The first step in discriminant function analysis was to ensure that the data met the assumptions of the analysis. One of the assumptions is the normal distribution of the values of each predictor (i.e. variable). Two commonly used statistical tests for normality of a given continuous variable are the Shapiro-Wilk and the Kolmogorov-Smirnov (K-S) tests (SPSS, 2007). However, the Shapiro-Wilk test is recommended for small to medium size samples less than 2000 observations (Garson, 2008), therefore the Kolmogorov-Smirnov test was used. With large sample sizes, unimportant deviations from normality may be significant (Garson, 2008), therefore frequency distributions and skewness significance tests are important to consider. Skewness significance tests were the same as used for the descriptive statistics in section 4.1.

In the preferred study area, the significance value from the K-S test indicated that the variables 38 kHz backscatter ($\alpha = 0.010$), depth ($\alpha = 0.000$), slope ($\alpha = 0.001$), and curvature were not normally distributed ($\alpha = 0.000$) and the 120 kHz backscatter ($\alpha = 0.060$) was normal, based on a 95% confidence interval. In the non-preferred study area, 38 kHz backscatter ($\alpha = 0.000$), depth less than 52m ($\alpha = 0.000$), depth greater than 52m ($\alpha = 0.000$), slope ($\alpha = 0.014$), and curvature were not normally distributed ($\alpha = 0.000$). However, the skewness significance tests in Table 29 indicated that the skewness values for all variables were not significant. In addition, the histograms in Figure 42 and Figure 43 indicated more or less normal distributions.

Table 29: Skewness significance test for discriminant function analysis input variables.

Variable	N	Skewness Value	Se	z	Significant
Non preferred Study Area					
38 kHz	9202	0.122	0.06	2.03	No
120 kHz	9889	0.122	0.06	2.03	No
Depth < 52m	1186	0.252	0.07	6.45	No
Depth > 52m	443	0.387	0.11	3.87	No
Slope	9207	-0.089	0.06	-1.45	No
Curvature	9889	0.142	0.06	2.37	No
Preferred Study Area					
38 kHz	9634	-0.236	0.06	-3.83	No
120 kHz	9534	0.030	0.06	0.50	No
Depth	9634	-1.023	0.06	-17.05	No
Slope	9535	0.005	0.06	0.13	No
Curvature	9634	-0.252	0.06	-4.20	No

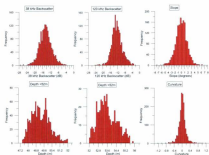


Figure 42: Histograms of discriminant function analysis input variables for the non-preferred study area.

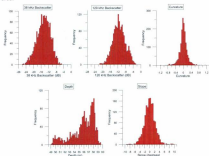


Figure 43: Histograms of discriminant function analysis input variables for the preferred study area.

The assumptions of linear relationships among all pairs of predictors within each group and homogeneity of variance-covariance matrices can be both tested using a matrix scatter plot (Figure 44 and Figure 45). The spreads of the scatter plots are roughly equal, therefore the assumption of homogeneity of variance-covariance matrices can be assumed (Leech et al., 2005). Absence of multicollinearity assumes there are no high intercorrelations among variables. The correlation matrix (Table 30) reveals the correlation values among the variables. Tabachnick and Fidell (2001) suggest that any correlation values higher than 0.7 would create issues.

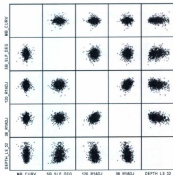


Figure 44: Matrix scatter plot of the input variables for discriminant function analysis in the non-preferred study area. Backscatter = 38_R1ADJ and 120_R1ADJ, Depth <52m = DEPTH_LS_52, Slope = SB_SLP_DEG, Curvature = MB_CURV. These plots aid in visualizing the relationship between predictors and the spread of each plot is roughly equal supporting the homogeneity of variance-covariance assumption.

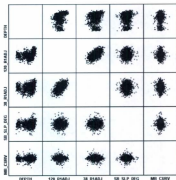


Figure 45: Matrix scatter plot of the input variables for discriminant function analysis in the preferred study area. Backscatter = 38_R1ADJ and 120_R1ADJ, Depth = DEPTH, Slope = SB_SLP_DEG, Curvature = MB_CURV. These plots aid in visualizing the relationship between predictors and the spread of each plot is roughly equal supporting the homogeneity of variance-covariance assumption.

Table 36: Correlation matrix of input variables for discriminant function analysis.

Preferred Study Area						
	38 kHz BS	120 kHz BS	Depth	Slope	Curvature	
38 kHz BS	1	0.557	0.299	0.026	-0.046	
120 kHz BS	0.557	1	0.293	0.026	0.096	
Depth	0.299	0.293	1	0.089	0.076	
Slope	0.026	0.026	0.089	1	-0.007	
Curvature	-0.046	0.096	0.076	-0.007	1	
Non-preferred Study Area						
	38 kHz BS	120 kHz BS	Depth<50m	Depth>50m	Slope	Curvature
38 kHz BS	1	0.305	-0.273	0.194	-0.071	-0.001
120 kHz BS	0.305	1	-0.201	0.026	0.036	-0.034
Depth<50m	-0.273	-0.201	1	-	0.176	-0.077
Depth>50m	0.194	0.026	-	1	0.134	-0.073
Slope	-0.071	0.036	0.176	0.134	1	-0.030
Curvature	-0.001	-0.034	-0.077	-0.073	-0.030	1

All of the variables were entered into the analysis simultaneously. Three tables provided in the output assess the contribution of each variable to the model: tests of equality of group means table (Table 31), discriminant function coefficients table (Table 32), and correlation between predictor variables and discriminant functions table (Table 33). The test of equality of group means table (Table 31) specifies each variable's potential to contribute to the model before the model is created with significance values greater than 0.10 indicating poor contribution to the model (SPSS, 2007). The Wilks' Lambda values are another indication of a variable's potential contribution with smaller values indicating better discrimination between the groups (i.e. ISUs). Both the curvature and the slope do not contribute well to the model in both study areas.

Table 31: Test of equality of group means of the discriminant function analysis for the preferred study area.

	Wilks' Lambda	F	df1	df2	Sig.
Preferred Study Area					
38 kHz BS	.987	31.115	6	1467	.000
120 kHz BS	.769	73.563	6	1467	.000
Slope	.999	.322	6	1467	.925
Curvature	.998	.456	6	1467	.945
Depth	.915	55.284	6	1467	.000
Non-preferred Study Area Depth <52m					
38 kHz BS	.849	52.985	4	1188	.000
120 kHz BS	.913	28.221	4	1188	.000
Slope	.997	.855	4	1188	.491
Curvature	.996	1.189	4	1188	.314
Depth<52m	.868	45.340	4	1188	.000
Non-preferred Study Area Depth >52m					
38 kHz BS	.747	25.743	6	456	.000
120 kHz BS	.705	32.518	6	456	.000
Slope	.993	.551	6	456	.766
Curvature	.994	.481	6	456	.822
Depth>52m	.865	12.352	6	456	.000

Discriminant function coefficients presented in Table 32 are standardized to allow comparison of variables measured on different scales. Larger absolute values correspond to variables with greater discriminating power (SPSS, 2007). For instance,

120 kHz backscatter and depth provide the best discriminating power for the first discriminant function in the preferred study area. Although there were five discriminant functions in the preferred study area, only the first two functions were significant (see the following paragraph concerning Wilks' Lambda and Table 34); the same applied to the non-preferred study area.

Table 32: Standardized discriminant function coefficients.

Variables	Function				
	1	2	3	4	5
Preferred Study Area					
38 kHz BS	.094	.964	.715	.149	-.241
120 kHz BS	.718	-.938	.183	-.018	.075
Slope	-.056	-.081	-.073	.893	.118
Curvature	-.036	.133	.294	-.085	.957
Depth	.545	.582	-.716	-.110	.093
Non-preferred Study Area Depth >52m					
38 kHz BS	.462	.119	.764	-.145	-.729
120 kHz BS	.678	-.250	-.604	.159	.674
Slope	-.191	-.170	.481	.843	.229
Curvature	.034	.085	.535	-.490	.696
Depth >52m	.119	.981	-.163	.981	.167
Non-preferred Study Area Depth <52m					
Depth <52m	-.572	.723	.214	.123	-
38 kHz BS	.587	.641	-.489	-.055	-
120 kHz BS	.347	.094	.935	.247	-
Slope	.126	-.458	.048	-.680	-
Curvature	.012	-.370	-.172	.728	-

Table 33 displays information about the correlation between each predictor variable and the discriminant scores produced by the functions. Higher values are an indication of the influence a certain variable has on a specific discriminant function score. For example, in the preferred study area, 120 kHz BS (0.850) and Depth (0.712) were highly correlated with the first function. 38 kHz BS (0.493), Depth (0.443) and 120 kHz BS (-0.336) were correlated with second function. Slope (0.984) was highly correlated with the third function, and so on.

Table 33: Correlation between predictor variables and discriminant functions.

Variables	Function				
	1	2	3	4	5
Preferred Study Area					
120 kHz BS	.850	-.336	.421	.053	-.024
Depth	.712	.443	-.533	.025	.109
38 kHz BS	.491	.493	.670	.163	-.218
Slope	.020	-.024	-.130	.984	.108
Curvature	-.014	.110	.205	-.090	.968
Non-preferred Study Area Depth <12m					
38 kHz BS	.738	.535	-.378	-.812	-
Depth<12m	-.662	.521	.202	-.836	-
120 kHz BS	.542	.109	.810	.171	-
Curvature	.004	-.364	-.209	.736	-
Slope	.018	-.306	.138	-.649	-
Non-preferred Study Depth >12m					
120 kHz BS	.894	-.206	-.202	.197	.281
38 kHz BS	.799	.042	.484	.035	-.380
Depth>12m	.125	.954	-.083	.240	.098
Slope	-.963	-.031	.454	.850	.259
Curvature	-.038	.024	.563	-.486	.887

The discriminant analysis procedure in SPSS also provided two tables that indicated how well the model as a whole fits the data, namely Wilks' Lambda (Table 34) and eigenvalues tables (Table 35) (SPSS, 2007). Wilks' Lambda measures how well each function separated cases into groups and is equal to the proportion of the total variance in the discriminant scores not explained by differences among the groups. Smaller values of Wilks' lambda indicate greater discriminatory ability of the function (SPSS, 2007), such as the first function in each study area. The Chi-square statistic in Table 34 tests the hypothesis that the means of the functions listed are equal across groups (i.e. ISUs) (SPSS, 2007). Small significance values for the first two functions in both study areas mean that they do better than chance at discriminating between ISUs; the remaining functions do not do better than chance at discriminating between ISUs. For instance, in the preferred study area, the first function explained 75.4% of the

variance, the second function explained 23.5% of the variance, and the remaining three functions only explained the remaining 1.1% of the variance (Table 35).

Table 34: Wilks' Lambda results for the preferred study area.

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
Preferred Study Area				
1 through 5	.624	668.109	30	.000
2 through 5	.865	175.840	20	.000
3 through 5	.904	6.230	12	.767
4 through 5	.999	1.614	6	.952
5	1.000	.091	2	.959
Non-preferred Study Area Depth <52m				
1 through 4	.738	360.524	20	.000
2 through 4	.968	39.167	12	.000
3 through 4	.995	5.889	6	.436
4	1.000	.099	2	.971
Non-preferred Study Area Depth >52m				
1 through 5	.957	271.804	30	.000
2 through 5	.840	79.248	20	.000
3 through 5	.983	7.833	12	.790
4 through 5	.996	1.855	6	.932
5	.999	.481	2	.786

Table 35: Eigenvalues and explained variance of the discriminant functions for the preferred study area.

Functions	Eigenvalue	% of Variance	Cumulative %
Preferred Study Area			
1	0.396	75.4	75.4
2	0.123	23.5	98.9
3	0.005	0.9	99.8
4	0.001	0.2	100.0
5	0.000	0.0	100.0
Non-preferred Study Area Depth < 52m			
1	0.211	90.5	90.5
2	0.028	9.3	99.8
3	0.005	1.4	100.0
4	0.0	0.0	100.0
Non-preferred Study Area Depth >52m			
1	0.500	73.8	73.8
2	0.189	23.8	97.6
3	0.013	1.9	99.5
4	0.003	0.4	99.9
5	0.001	0.1	100.0

Mean discriminant scores were also calculated for each ISU based on unstandardized discriminant functions (Table 36). In the preferred study area, the first function performed rather well in discriminating between the ISUs with the exception of

GRS and SM that had similar values. The second function did a particularly poor job discriminating between SM and SG, as well as GRI and GL. In the non-preferred study area with depth <52m, the first function displayed little difference between mean discriminant scores of Sand and GRI. The second function showed little difference between Sand, GRI, and SM. In the non-preferred study area with depth >52m, the first function displayed little difference between mean discriminant scores of SM and SB, and to a lesser extent GRS and Sand. The second function showed little difference between GRI and GH. Scatter plots of the scores of each observation for the first two functions, coloured coded by the ISU for which each case belongs, provided useful information on the distribution of discriminant function scores in multivariate space (Figure 46, Figure 47, and Figure 48).

Table 36: Mean discriminant scores for each function.

ISU	Function				
	1	2	3	4	5
Preferred Study Area					
GRS	-420	-649	010	-012	-008
GRI	920	-176	302	827	807
GR	-667	385	094	893	-017
GL	2008	458	201	-112	-024
Sand	1527	129	404	898	824
SM	-324	223	-081	-013	804
SG	610	221	-189	808	-014
Non-preferred Study Area Depth <52m					
GRS	-3581	1244	-082	122	-
GR	-721	087	-243	-087	-
GL	-1021	3548	499	-068	-
Sand	-487	-073	341	-081	-
SM	528	033	-006	801	-
Non-preferred Study Area Depth >52m					
GRS	-1244	-623	-111	828	826
GRI	886	112	187	171	-015
GL	956	-279	-034	-015	012
GR	-182	241	334	015	-018
Sand	-1456	-1298	827	-182	826
SM	-346	2091	192	129	347
SB	-415	847	-082	-130	006

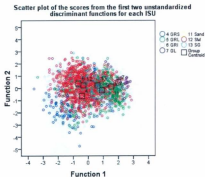


Figure 46: Scatter plot of the scores for the first two discriminant functions with corresponding ISUs for the preferred study area.

Scatter plot of the scores from the first two unstandardized discriminant functions for each ISU

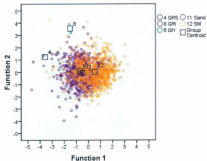


Figure 47: Scatter plot of the scores for the first two discriminant functions with corresponding ISUs for the non-preferred study area with depth <52m.

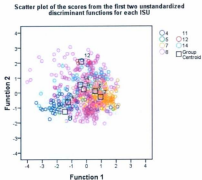


Figure 48: Scatter plot of the scores for the first two discriminant functions with corresponding ISUs for the non-preferred study area with depth >52m.

As a final step, discriminant function analysis provided classification functions that assign each observation of the input variables to an ISU. Each ISU had its own classification equation with classification coefficients for each variable (Table 37 and Table 38).

Table 37: Classification function coefficients for the preferred study area.

Variables	ISUs						
	GRS	GRL	GRI	GL	Sand	SM	SG
38 kHz BS	-2.101	-1.974	-1.804	-1.788	-1.795	-1.872	-1.911
120 kHz BS	-3.326	-3.148	-3.717	-3.071	-3.095	-3.585	-3.387
Slope	-1.903	-1.836	-1.890	-2.059	-1.944	-1.937	-1.947
Curvature	-8.493	-8.365	-7.727	-7.996	-7.637	-7.886	-8.414
Depth	24.915	25.630	25.111	26.024	25.628	25.231	25.641
(Constant)	-729.636	-760.813	-741.063	-786.498	-763.868	-747.953	-788.414

Table 38: Classification function coefficients for the non-preferred study area.

Variables	ISUs							
	GRS	GRL	GRI	GL	GH	Sand	SM	SB
Non-preferred Study Area with Depth <52m								
Depth<52m	72.672	-	68.636	-	73.244	68.433	68.839	-
38 kHz BS	.987	-	1.426	-	2.041	1.375	1.688	-
120 kHz BS	-.684	-	-.299	-	-.003	-.130	.836	-
Slope	-7.484	-	-6.927	-	-7.892	-6.864	-6.821	-
Curvature	-.378	-	1.960	-	-8.143	2.076	1.955	-
(Constant)	-1865.57	-	-1722.87	-	-1896.67	-1711.50	-1675.23	-
Non-preferred Study Area with Depth >52m								
	GRS	GRL	GRI	GL	GH	Sand	SM	SB
38 kHz BS	-2.622	-2.302	-	-2.311	-2.423	-2.445	-2.426	-2.466
120 kHz BS	.584	.789	-	.907	.622	.310	.498	.571
Slope	-7.096	-7.252	-	-7.425	-7.388	-6.791	-7.327	-7.430
Curvature	20.680	21.413	-	21.309	21.313	22.906	22.989	21.485
Depth>52m	82.075	80.236	-	82.824	83.294	80.996	85.670	83.682
(Constant)	-2402.32	-2515.60	-	-2493.58	-2522.84	-2405.93	-2653.81	-2543.85

In the preferred study area, only 26.1% of the observations were correctly classified with this analysis (Table 42). In the non-preferred study area, 50.6% was correctly classified for the depth<52m dataset and 34.1% was correctly classified with depth>52m (Table 42). There were patterns with regards to misclassification of some of the ISUs. In the preferred study area (Table 38), the model does particularly well classifying GL (70.6%), Sand (75%), and to a lesser extent, GRI (55.7%). However, there were low classification accuracies for SM (14.8%), GRL (20.9%), and SG (28.0%). Most of the SM samples were misclassified as GRI (38.6%) and GRS (22.2%). SG

(19.5%), GL (12.5%), and GRS (12.5%) were also common misclassifications for GRL. The SG ISU was commonly misclassified as either GRL (20.4%) or Sand (18.3%).

In the non-prefereed study area with depth < 52m (Table 40) the model classified GRS (100.0%) and GH (100.0%) perfectly, however there were only 3 and 2 observations respectively. SM (55.7%) had the best classification accuracy, followed by Sand (35.6%) and GRI (29.3%). GRI was commonly misclassified as SM (26.7%) and Sand (25.3%). In the non-prefereed study area with depth > 52m (Table 41) the model classified Sand and SM with 100% accuracy, but again there were only a few observations for each ISU. The model classified GL (65.9) well, and GRS (58.6) was also classified relatively well. GH was classified rather poorly with 6.5% accuracy, and neither of the ISUs seemed to dominate the misclassification. SB was only classified with 22.2% accuracy, with 25% of its observations misclassified as GRL and 19.6% misclassified as GRS.

Table 38: Classification results of the discriminant function analysis for the preferred study area.

	ISUs	Predicted Group Membership							Total
		GRS	GRL	GRI	GL	Sand	SM	SG	
Original Count	GRS	122	34	36	4	10	19	44	269
	GRL	36	80	20	36	72	7	66	287
	GRI	13	4	44	0	0	15	3	79
	GL	0	2	0	24	4	0	4	34
	Sand	0	1	0	0	3	0	0	4
	SM	157	46	273	25	36	165	60	708
	SG	2	19	14	3	17	12	26	83
%		GRS	GRL	GRI	GL	Sand	SM	SG	
	GRS	45.4	12.6	13.4	1.5	3.7	7.1	16.4	100.0
	GRL	12.5	28.9	7.0	12.5	25.1	2.4	19.5	100.0
	GRI	16.5	5.1	55.7	0	0	19.0	3.8	100.0
	GL	0	5.9	0	70.6	11.8	0	11.8	100.0
	Sand	0	25.0	0	0	75.0	0	0	100.0
	SM	22.2	6.5	38.6	3.5	5.1	14.8	9.3	100.0
	SG	2.2	20.4	18.1	3.2	18.3	12.9	28.8	100.0

Table 40: Classification results of the discriminant function analysis for the non-preferred study area with depth <52m.

	ISUs	Predicted Group Membership					
		GRS	GRL	GL	Sand	SM	Total
Original Count	GRS	3	0	0	0	0	3
	GRL	7	22	7	19	20	75
	GL	0	0	3	0	0	3
	Sand	13	152	4	183	132	584
	SM	6	86	11	103	394	600
%		GRS	GRL	GL	Sand	SM	
	GRS	100.0	.0	.0	.0	.0	100.0
	GRL	9.3	29.3	9.3	25.3	36.7	100.0
	GL	.0	.0	100.0	.0	.0	100.0
	Sand	2.5	35.4	.8	35.6	25.7	100.0
	SM	1.0	14.3	1.8	17.2	65.7	100.0

Table 41: Classification results of the discriminant function analysis for the non-preferred study area with depth >52m.

	ISUs	Predicted Group Membership						
		GRS	GRL	GL	GH	Sand	SM	Total
Original Count	GRS	34	3	5	2	11	2	58
	GRL	1	7	3	2	0	0	14
	GL	4	29	87	9	1	0	132
	GH	31	30	46	14	21	44	215
	Sand	0	0	0	0	5	0	5
	SM	0	0	0	0	0	3	3
	SB	7	9	2	2	4	4	30
		GRS	GRL	GL	GH	Sand	SM	SB
%	GRS	58.6	5.2	8.6	3.4	19.0	3.4	100.0
	GRL	7.1	50.0	21.4	14.3	.0	.0	100.0
	GL	3.0	22.0	65.9	6.8	.8	.0	100.0
	GH	14.4	14.0	21.4	6.8	9.8	20.9	100.0
	Sand	.0	.0	.0	.0	100.0	.0	100.0
	SM	.0	.0	.0	.0	.0	100.0	100.0
	SB	18.4	25.0	5.8	5.6	11.1	11.1	100.0

In order to determine if using backscatter values from both frequencies improved classification accuracy, the supervised classification was run using backscatter values from each frequency independently, and then both frequencies together, along with the morphology layers. The 120 kHz frequency did slightly better than the 38 kHz frequency, with the exception of the non-preferred study area with depth <52m (Table 42). Although the difference was small, there was slight improvement in the classification of ISUs when

using backscatter from two frequencies as opposed to one (Table 42). However, these classification accuracy values are average classification accuracy values among all the ISUs and it must be kept in mind the accuracy values vary greatly among the ISUs (c.f. Table 39, Table 40, and Table 41).

Table 42: Classification accuracies for the supervised classification results.

Frequencies Used	Classification Accuracy
Preferred Study Area	
38 kHz and 120 kHz	35.1%
38 kHz	21.0%
120 kHz	34.1%
Non-preferred Study Area depth<52m	
38 kHz and 120 kHz	50.6%
38 kHz	49.5%
120 kHz	45.1%
Non-preferred Study Area depth>52m	
38 kHz and 120 kHz	34.1%
38 kHz	30.3%
120 kHz	32.2%

4.5.2 Unsupervised k-means Clustering

The k-means output in SPSS has an ANOVA table that provides an indication of which variables contributed the most to the cluster solution. The relative size of the statistics provides information about each variable's contribution to the separation of the groups, with larger *F* values indicating greater separation between clusters (SPSS, 2007). However, the *F* tests should only be used for descriptive purposes and the significance levels cannot be interpreted as tests of the hypothesis that the cluster means are equal (SPSS, 2007). In the preferred study area, depth contributed most to the separation of the clusters (2195.76) followed by 38 kHz (517.04) and 120 kHz backscatter (471.97), slope (37.39), and curvature (0.83). In the non-preferred study with depth <52m, depth contributed the most to the separation of clusters (2457.20) followed by 38 kHz backscatter (56.34), slope (55.14), 120 kHz backscatter (32.99), and curvature (0.78). In the non-preferred study area with depth >52m depth contributed

most to cluster separation (296.88) followed by 120 kHz backscatter (208.09), 38 kHz backscatter (191.79), slope (7.13), and curvature (6.70).

The Euclidean distances between final cluster centroids in multivariate space conveyed the level of similarity, or dissimilarity, between the clusters (Table 43, Table 44, and Table 45). These tables revealed how far, and in a sense how different, clusters were from one another. In other words, greater distances correspond to greater dissimilarities (SPSS, 2007). These distance values were based on the normalized 0 to 1 scale used for the input data. For instance, the distance between class 1 and class 2 in the preferred study area is large (0.657) relative to the distance between class 1 and class 3 (0.248). Results of both k-means classifications were overlaid on the interpreted surficial geology and are displayed in Figure 49, Figure 50, and Figure 51. At first glance, it was not obvious that any particular class from the k-means clustering directly related to any specific seabed type. In the non-preferred study area with depth <52m, the k-means classes seemed to be influenced by the trend of the depth. In the same study area with depth >52m, the spatial patterns of the classes seemed more or less random. However, one could argue that class 3 is correlated with GRL. As for the preferred study area, class 5 seemed to mainly overlay GRL and GL and class 3 commonly occurred over SM.

Table 43: Distances between final cluster centers for the k-means classification in the preferred study area.

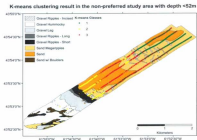
Cluster	1	2	3	4	5	6
1		.657	.248	.272	.302	.451
2	.657		.447	.638	.773	.403
3	.248	.447		.349	.365	.367
4	.272	.638	.349		.541	.289
5	.302	.773	.365	.541		.687
6	.451	.403	.367	.289	.687	

Table 44: Distances between final cluster centers for the k-means classification in the preferred study area with depth <52m.

Cluster	1	2	3
1		.513	.247
2	.513		.269
3	.247	.269	

Table 45: Distances between final cluster centers for the k-means classification in the preferred study area with depth >52m.

Cluster	1	2	3	4	5
1		.550	.524	.261	.297
2	.550		.568	.405	.391
3	.524	.568		.321	.664
4	.261	.405	.321		.369
5	.297	.391	.664	.369	

**Figure 49:** Map of the clustering results overlaid on the interpreted surficial geology in the non-preferred study area with depth <52m.

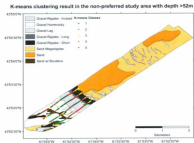


Figure 50: Map of the clustering results overlaid on the interpreted surficial geology in the non-preferred study area with depth >52m.

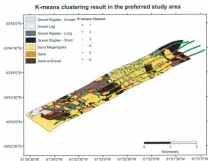


Figure 51: Map of the clustering results overlaid on the interpreted surficial geology.

The clusters from the k-means clustering with the highest frequency of occurrence (i.e. the mode) for each ISU are outlined in Table 46. In the preferred study area, the mode for GRS was class 4, but this ISU also had a high occurrence of class 1 and 6. Similarly, SG had a mode of class 1 but had a high occurrence of class 5. The remaining ISUs had quite distinct modes. For the non-preferred study with depth < 52m, GRI had a bimodal distribution with classes 2 and 3, while the remaining ISUs had distinct modes. In the non-preferred study area with depth > 52m, all ISUs had distinct modes with the exception of GL which had a high occurrence of class 1 along with class 5. There are ISUs in each study area that have the same k-means class mode, meaning that the values of the input variables for these ISUs were very similar for specific clusters.

Table 46: k-means class modes for each ISU in both study areas.

ISU	Preferred Study Area	Non-preferred Study Area with depth < 52m	Non-preferred Study Area with depth > 52m
GRS	4	-	3
GRL	5	-	5
GRI	3	2 & 3	-
GL	5	-	5
GPI	-	-	2
Sand	-	3	-
SM	3	1	-
SD	1	-	-
SG	-	-	5

4.6 Summary of Results

Table 47 and Table 48 provides an overall summary of the main results from chapter 4. The information in Table 47 reiterates a number of general trends from the mean near-nadir backscatter values observed in this chapter. 120 kHz mean backscatter was higher than 38 kHz mean backscatter, gravel seabeds had higher mean backscatter than sand seabeds, and mean backscatter values in the preferred area were generally higher than the non-preferred study area.

Table 47: Summary of mean near-nadir backscatter values within each study area summarized by frequency and dominant substrate type.

	Non-preferred Study Area	Preferred Study Area
Mean BS 120 kHz	-13.43	-10.79
Mean BS 120 kHz Gravel Seabeds	-11.91	-10.37
Mean BS 120 kHz Sand Seabeds	-12.30	-12.00
Mean BS 38 kHz	-14.86	-13.79
Mean BS 38 kHz Gravel Seabeds	-14.11	-14.67
Mean BS 38 kHz Sand Seabeds	-14.80	-14.69

Table 48 summarizes the main results observed from the individual ISUs. In sum, GRS, GRI, SB, and SG generally did not follow typical trends of gravel and sand substrate seabeds based on mean near-nadir backscatter values, quantile classifications, and frequency differences based on quantile classification differences. However, SG only seemed to deviate from typical sand substrate trends based on the 120 kHz frequency. These ISU also had relatively low classification accuracies with the exception of GRS in both study areas and GRI in the preferred study area only.

Table 48: Summary of general results of each ISU. QC = Quantile classification results, QCD = quantile classification differencing results, DFA (%) = supervised classification accuracy percent. QC and QCD values (Y = yes and N = no) indicate whether the ISU followed the typical trend of higher near-nadir backscatter from gravel substrate then sand substrate. The Wilcoxon values indicate whether the mean near-nadir backscatter values from each frequency was significantly different for each ISU.

	GRS	GRI	QC	SH	Sand	SB	SG	SB
Non-preferred Study Area								
Mean BS 120 kHz	-13.90	-8.90	-14.07	-7.63	-11.30	-13.49	-10.41	-10.56
Mean BS 38 kHz	-12.72	-13.43	-15.03	-11.30	-14.34	-15.79	-13.81	-10.39
Variance 120 kHz	13.81	8.28	7.53	5.67	13.69	3.25	4.64	14.06
Variance 38 kHz	13.29	11.56	7.78	7.63	12.19	4.45	5.36	16.68
QC 120 kHz	N	Y	N	Y	Y	N	-	N
QC 38 kHz	N	Y	N	Y	Y	N	-	N
QCD	Y	Y	N	Y	Y	Y	-	N
Wilcoxon	Y	N	Y	Y	N	N	Y	-
DFA %	88.6	90.8	29.3	95.9	8.5	38.9	65.7	22.2
Preferred Study Area								
Mean BS 120 kHz	-11.71	-8.94	-13.13	-6.53	-	-7.82	-10.30	-10.24
Mean BS 38 kHz	-14.40	-13.40	-14.74	-10.54	-	-11.50	-14.85	-13.84
Variance 120 kHz	7.86	8.38	8.18	5.96	-	1.76	8.60	12.29
Variance 38 kHz	14.46	10.42	10.94	10.90	-	13.02	11.28	10.28
QC 120 kHz	N	Y	N	Y	-	Y	Y	-
QC 38 kHz	N	Y	N	Y	-	Y	Y	-
QCD	Y	Y	N	Y	-	N	Y	N
Wilcoxon	Y	Y	Y	N	-	N	Y	N
DFA %	45.4	20.9	95.7	70.8	-	75.9	14.8	28.0

5. Discussion

This section summarizes and discusses the main results from the analyses and relates them to other relevant studies. Five main discussion sections are outlined following the sections of the results chapter and how these results are used to answer the research questions. These sections are titled as follows: 5.1 Univariate Acoustic Backscatter Analysis, 5.2 Univariate Acoustic Backscatter Classification, 5.3 Frequency Differences, 5.4 Multivariate Acoustic Backscatter Analysis, and 5.5 Multivariate Classification. Finally, some important issues surrounding the accuracy of both the interpreted surficial geology map and the acoustic data are discussed in section 5.6 General Considerations.

5.1 Univariate Acoustic Backscatter Analysis

The results of the statistical analysis of backscatter values from each frequency (see section 4.1) answered research question 1, which stated: "what are the univariate (single frequency) statistical characteristics of near-nadir acoustic backscatter data from each frequency independently?" Answering this question provides a comparison of how near-nadir backscatter from each frequency responded to the seabed. The results revealed trends that were expected based on information from the literature. The 120 kHz data generally had higher near-nadir backscatter values than the 38 kHz data (Table 7 and Table 8). This was expected as previous empirical studies of multi-frequency acoustic seabed backscatter have revealed that backscatter values increased with increasing frequency (McKinney and Anderson, 1964, Jackson et al., 1986, Jackson and Richardson, 2007). This could be explained by the shorter wavelengths associated with a higher frequency exhibiting less penetration and absorption of sound energy into the

seabed, which results in a dominant surface scattering process that is highly correlated with sediment size (Galloway and Collins, 1998, Kloser et al. 2001, Chakraborty et al. 2007). Courtney et al. (2005) who also used the 120 kHz dataset to classify substrate in this study area stated that the coarse grain sizes results in little penetration of the 120 kHz frequency. This reflects the relationship between wavelength and sediment size, where backscatter generally decreases as the dimensions of a target (i.e. grain size of seabed substrate) decrease relative to the acoustic wavelength (Lurton, 2001). However, it must be mentioned that there are studies that reveal near-nadir backscatter is higher at lower frequencies for certain substrate types based on the models used (APL94, 1994). One could hypothesize that the seabed may "appear" rougher to higher frequencies at near-nadir due to its smaller wavelength causing the signal to scatter away from the sensor, causing the intensity of the signal to decrease and elongating the signal and shape of the return.

Another expected result confirmed by the descriptive statistics analysis was that gravel seabeds had significantly higher near-nadir backscatter values than sand seabeds with the exception of the 38 kHz frequency in the preferred study area (Table 8). Gravel substrate has a larger grain size and is generally denser than sand substrate, which leads to higher backscatter values due to less penetration and absorption of sound waves and higher impedance values (Stemlich and de Moustier, 2003, Manik et al., 2006). The lack of a significant difference between sand and gravel seabeds for the 38 kHz frequency in the preferred study area could be due to two reasons. First, perhaps the high degree of patchiness, as defined by the relative size, shape, and density of the distribution of ISUs (Fader, 2007) was more variable in this study area causing acoustic backscatter to be more variable and therefore making it difficult to differentiate homogenous areas of sand and gravel. Second, results from this, and previous

(Chakraborty et al., 2007), studies suggest that a lower frequency (38 kHz) is less correlated with sediment size compared to a higher frequency (120 kHz). These two factors combined provide an explanation for this result.

Average near-nadir backscatter values from both frequencies were also significantly higher for the preferred study area (Table 7). This is likely due to the higher amount of gravel in the preferred study area (40.3%) as opposed to the non-preferred study area (34%). The difference between frequencies in terms of average backscatter values was larger for gravel seabeds compared to sand seabeds (Table 8). There was little difference between mean 38 kHz backscatter values of sand and gravel seabeds, suggesting that this frequency may be less effective for discriminating between the two substrate types. This result is consistent with those of Chakraborty et al. (2007), where higher frequency (210 kHz) backscatter were believed to have a dominant water-seabed surface scattering and lower frequency backscatter have a dominant substrate volume backscatter. As their results reveal, 210 kHz frequency backscatter proved more effective in discriminating seabed types, although there was still important information embedded in low frequency backscatter coming from within the volume of the substrate. However, there certainly is a limit to the highest frequency to use for discriminating substrate type, as wavelengths may become too short to detect sediment.

In terms of average near-nadir backscatter values associated with each interpreted sediment unit (ISU) from the surficial geology maps, gravel ripples – short wavelength (GRS) and gravel ripples incised (GRI) values had low backscatter values more similar to sand seabeds (Table 9). Perhaps this is because the backscatter signal was influenced by the roughness of these seabeds (i.e. gravel ripples) causing some of the signal to be scattered away from the sensor and decreasing the amount of specular reflection (Lorton, 2002). The texture and roughness of these seabeds could have

possibly been embedded in acoustic backscatter signals as their spatial extents were within the size of the acoustic footprint. GRS seabeds have ripples of 2-3 meter wavelength size and GRI seabeds are "ribbons" of incised gravel ripples within sand seabeds and have typical width values within the extent of the acoustic footprint. Sand to sandy gravel (SG) also had similar mean backscatter values to gravel seabeds (Table 9). This could be a result of the sound waves detecting the gravel sediments within this seabed type and therefore increasing the backscatter intensity. The variability in the backscatter of these seabeds could also be attributed to the transitions between different seabed types and how this could create variable backscatter responses.

At least one acoustic frequency had a difference greater than 1 between mean near-nadir backscatter values when comparing one ISU to another (Table 10 and Table 11). This suggests that using two frequencies, as opposed to one, should help to better discriminate between different surficial seabed types based on near-nadir backscatter values. However, certain pairs of ISUs (SM and GRI in the preferred study area; Sand and GRI, and Sand and SB in the non-preferred study area) did not show much variation between mean backscatter values for either frequency. Hence, it could be concluded that these particular pairs of ISUs may be difficult to differentiate.

Certain ISUs (SG, SB, GH, and GRS) had high variance values indicating that the near-nadir acoustic backscatter was more variable in reference to the mean within each ISU (Table 9). This is likely due to the heterogeneous nature of these seabeds resulting from a mixture of grain sizes as well as high degrees of roughness. This assumption is based on findings by Collier and Brown (2005), who revealed that a heterogeneous mix of sediments produced more variable backscatter intensity measurements than homogenous substrate. SG had high variance values for the 120 kHz frequency and low for the 38 kHz frequency (Table 9). Perhaps the 120 kHz

frequency was able to detect the mixture of sediments in this class, as it is more sensitive to sediment size (Gloser et al., 2001), resulting in higher variance values. ISUs with high variance values are expected to have lower classification accuracy than those with low variance values, because typical backscatter responses for these ISUs will be harder to define (see section 5.5). Finally, the high variance values could be a result of transition zones between the ISUs causing variation on backscatter responses.

5.2 Univariate Acoustic Backscatter Classification

The results of single frequency acoustic backscatter classification answered research question 2, which stated the question: "how do univariate classifications of acoustic backscatter values relate to seabed substrate types?" Addressing this question provides a basis to compare single frequency and dual-frequency results. Contrary to expectation, 38 kHz quantile classification results revealed that sand seabeds in the non-preferred study area had more occurrences of high quantile classes synonymous to higher backscatter values (Table 15) and gravel seabeds in the preferred study area had more occurrences of low quantile classes synonymous to lower backscatter values (Table 14). However, as noted in section 4.2, these results were influenced by specific ISUs (SM for the sand seabeds and GRS and GRI for the gravel seabeds). Therefore, one must be cautious when generalizing the study areas into sand and gravel seabeds, as there is variability within each ISU. Similar to the descriptive statistics results, perhaps these unexpected results can be explained by the following: transitions between seabeds; the various morphological features associated with each seabed type; heterogeneous mixture of sediments.

These initial classification attempts using each frequency independently (i.e. single frequency classification) proved that it was difficult to relate results to specific

spatial features from the interpreted surficial geology classifications. At most, such single frequency classifications of near-nadir backscatter values may discriminate between distinct sand and gravel substrate areas based on high and low backscatter values, with the 120 kHz frequency providing better discrimination. These results were expected, as previous studies that used a single frequency to define seabed substrate types in this study area could only attribute classification results of acoustic backscatter to sand and gravel classes (Courtney et al., 2005).

5.3 Frequency Differences

Results from section 4.3 answered research question 3, which stated: "where do the major frequency differences occur and what are the spatial scales at which they occur?" The first attempt at understanding variations in local acoustic frequency responses was the differencing of the quantile classes representing a measure of the magnitude of difference between classified near-nadir backscatter values from both frequencies (see section 4.3.1). Larger differences were interpreted as locations where differences between quantile classes of the two frequencies were greater than the rest of the study area (Table 16 and Figure 21). The sand seabeds in the non-preferred study area had the highest occurrence of frequency differences based on classified values (Table 17). Classified backscatter values from the 120 kHz frequency were relatively higher than the 38 kHz frequency for gravel seabed types, while the opposite was observed for sand seabed types (Table 17). Therefore, these results indicate that there is a frequency dependent response from these two main substrate types. However, knowing what was driving this difference is difficult to determine.

SB and SG followed an opposite pattern than the one typically observed, with the quantile classification differencing values indicating that higher 120 kHz frequency

quantile classes occurred more than the 38 kHz frequency (Table 18 and Table 19). GRI had just the opposite with more occurrences of higher quantile classes from the 38 kHz frequency (Table 18 and Table 19). These ISUs also did not follow the observed patterns from the descriptive statistics for typical sand and gravel seabed (see section 5.1). SB is a mixture of sand and gravel within a boulder field surrounded by gravel seabed types, and SG is a mixture of sand and gravel. Perhaps there were more occurrences of higher classified backscatter values from the 120 kHz because this frequency was detecting the mixture of sediments better than using the 38 kHz frequency. In addition, GL had a very low occurrence of differences in classified backscatter values. This ISU is an example of a relatively flat and homogeneous seabed and therefore there is little variation for the frequencies to detect.

Frequency differences were also analyzed using continuous (i.e. unclassified) backscatter values. It was evident from the line graphs of dual-frequency near-nadir backscatter (Figure 26 and Figure 27) that at the scale of hundreds to thousands of meters, backscatter from both frequencies followed a general pattern of low (synonymous to sand seabeds) and high (synonymous to gravel seabeds) backscatter areas. Frequency differences of backscatter values were identified by comparing the slope of linear regression lines of both frequencies at a scale of tens to hundreds of meters (see section 4.3.2). The scale at which frequency divergences were examined proved to affect the results. The percentage of divergences for sand seabeds was higher than gravel seabeds in the non-preferred study area, and there was little difference between the percentage of divergences between sand and gravel seabeds in the preferred study area (Table 23). There were significantly less frequency divergences for gravel seabeds in the non-preferred study area compared to the preferred study (Table 23). The distribution of gravel in the non-preferred study area is all in the south-west part

of the map, with the exception of GRI. In contrast, the preferred study area has a higher percentage of gravel ISUs that are more dispersed among the sand ISUs. Perhaps the greater mix of sand and gravel ISUs caused more frequency divergences in the preferred study area.

The Wilcoxon Matched-pairs Signed-ranks test revealed significant differences in near-nadir backscatter values from the two frequencies over GRS, GRI, GL, and SM in the non-preferred study area (Table 24) and GRS, GRL, GRI, SM, and GL in the preferred study area (Table 25). SB, SG, and GH showed no significant differences in backscatter from the two frequencies. This is presumably due to the high variance values of BS values for these ISUs, meaning there was no consistent difference in backscatter values from the two frequencies due to high variability.

5.4 Multivariate Acoustic Backscatter Analysis

Results from the multivariate acoustic backscatter analysis answered research question 4, which stated the question: "what are the multivariate statistical characteristics of acoustic backscatter from both frequencies?" Correlation and regression analysis were used to here to provide insight into the relationship between near-nadir backscatter from the two frequencies. Correlation between the two frequencies resulted in values $R = 0.56$ for the preferred study area and $R = 0.52$ for the non-preferred study area, which are not particularly strong correlation values. This is a positive result as it indicates that the frequencies produced different near-nadir backscatter values in certain areas, which may provide additional information. The low r^2 values ($\sim 30\%$) obtained from the regression analysis further suggested that variations in one frequency were limited in helping to explain variations in the other frequency and there are significant portions of the seabed where the acoustic response from the two frequencies differs (see section

4.4). Hence, the remaining 70% of unexplained variance is either due to differences in frequency response or acoustic noise.

GRS, GRL, and GL in the non-preferred study area and GL, GRS, and GRI in the preferred study area had high residual values (Table 26). Generally, GRS and GRI also had high variance values (Table 9) and quantile class distributions that did not fit the expected trend for gravel seabeds (Table 18 and Table 19). The preferred study area had higher average residual values (Table 26), which probably reflects the spatial complexity of the different seabed classes in this area, as the acoustic footprints are more likely to overlap more than one ISU, which could create higher degrees of variability.

The low r^2 values from the linear regression models of the 38 kHz residuals and the morphology variables could be attributed to the fact that the morphology variables measure variations in the morphology of the seabed (i.e. depth, slope, curvature) that could vary across tens or hundreds of meters across the bank. On the other hand, backscatter only represents variations in substrate properties within the extent of the acoustic footprint. These results could also be attributed to noise in the acoustic signal.

5.5 Multivariate Classification

Both research question 5 and 6 were answered by the results of the multivariate classifications. Research question 5, "how do multivariate classifications of dual-frequency acoustic backscatter values relate to seabed substrate types", was addressed by comparing classified results to the ISUs from the interpreted surficial geology maps. Research question 6, "does a comparison of dual-frequency acoustic backscatter improve the ability to classify seabed substrate types in comparison to a single

frequency", was answered by assessing the accuracy of supervised classifications of single and dual-frequency results as discussed in the following paragraphs.

Unsupervised k-means clustering was used as an exploratory statistical method because the number of clusters was not based on statistically significant clusters but rather based on the number of ISUs within each study area. Cluster separation seemed to be mainly dependent upon depth, followed by backscatter (see section 4.5.2). It was difficult, however, to associate clusters to the certain ISUs.

The most obvious trend in the discriminant function classification results was the low classification accuracy values for the preferred study area (Table 42). The preferred study area has a much more complex spatial arrangement of ISUs, with a larger number of individual ISU polygons on the map (206) compared to the non-preferred study area (103). Therefore, one could reasonably expect more difficulty in classifying the preferred study area as there are more frequent changes in seabed classes. This same reasoning could also apply to the non-preferred study area with depth >52m. In terms of supervised classification accuracies for each frequency, the 120 kHz frequency performed better than the 38 kHz, with the exception of the non-preferred study area with depth <52m (Table 42). These results confirm findings of Chakraborty et al. (2007) who determined that a higher frequency (210 kHz) provide better discrimination between sediment types compared to a lower frequency (33 kHz). More importantly, these classification results revealed that near-nadir backscatter from two frequencies improves slightly upon classification with backscatter from one frequency (Table 42). These results are consistent with a number of other research papers that conclude that the use of multiple frequencies improves classification accuracy as multi-frequency backscatter provide additional information (Koser et al., 2002, Regi and Purkis, 2005).

SB, SG, GRI (non-preferred study area only) and GRS (preferred study area only) had low (<50%) classification accuracies (Table 39, Table 40, and Table 41) and high variance values (Table 8) as well as having gone against the expected trends of typical near-nadir backscatter values for sand and gravel seabeds (Table 18 and Table 19). SM and GRI in the preferred study area and GH and Sand in the non-preferred study area also had low classification accuracies. Higher degrees of roughness and mixed sediments associated with some of these seabeds may have caused high variance and inconsistent near-nadir backscatter responses. Therefore, the classification algorithm may have not effectively identified typical backscatter values for these seabeds, which the algorithm used to discriminate between the ISUs. In addition, GH, SB, and Sand in the non-preferred study area and SG in the preferred study area had no significant difference between dual-frequency backscatter from the Wilcoxon test results (Table 24 and Table 25) which implies that these ISUs may be difficult to differentiate based on near-nadir backscatter values. Finally, certain ISUs with low classification accuracies had been misclassified as another ISU that had similar mean near-nadir backscatter values from the descriptive statistics. For instance, SM in the preferred study area was commonly misclassified as GRI, and the differences in mean backscatter values of the two ISUs was 0.83 for the 120 kHz and 0.11 for the 38 kHz. GRI in the non-preferred study area was commonly misclassified as Sand, and the differences in the mean backscatter values of these two ISUs was also rather low (0.58 for the 120 kHz, and 0.16 for the 38 kHz).

5.6 General Considerations

It is important to keep in mind when discussing these results that the interpreted sediment units (ISUs) result from expert interpretations (Fader, 2007) and are hence only

a proxy of actual physical seabed characteristics in the natural environment for which accuracy is unknown. Sharp boundaries between ISUs are an artefact of classification for the purposes of human interpretation, whereas in reality transitions between physical seabed types can be more gradual. Nevertheless, the ISUs identified on the surficial geology maps were based on data collected at a much higher resolution (e.g. 0.25 m spatial resolution of sidescan imagery) compared to the acoustic footprint sizes (smallest footprint sizes being 8.5 m x 14.5 m) of the SBES data used to classify the seabed in this study. Therefore, it cannot be expected that the SBES backscatter data will detect all the finer details of seabed features that the sidescan data could. In addition, acoustic observations taken at or near the boundaries between ISUs resulted in acoustic footprints encompassing more than one ISU, whereas in the data, each observation was assigned to only one ISU. Therefore, it is likely that samples that occurred close to ISU boundaries are more variable in terms of backscatter. Another spatial issue related to the acoustic data is that the dGPS accuracy was ± 3 m, meaning at best the spatial location of a footprint was only within 6 m of the true position. It is also important to consider that the acoustic footprints penetrated within the seabed resulting in volume scattering and therefore the resulting near-nadir backscatter response could of been influenced by the substrate's volume, which may not be reflected in the ISU interpretations. The penetration capability also differed between frequencies, ranging from centimetres for the 120 kHz and tens of centimetres for the 38 kHz based on the discussions of Galloway and Collins (1996). Finally, another issue faced is that acoustic data are, by nature, noisy. Data points were averaged over five pings to reduce noise, but acoustic backscatter is still very complex and has a low signal-to-noise ratio.

There are also two considerations related to footprint differences between frequencies. First, the footprint size was slightly different between frequencies because

of the different angles used to define the near-nadir backscatter component (see section 3.3). As a result, backscatter from each frequency represented slightly different sized areas ranging from 1 to 2 m in both width and length (approximately 30% percent difference). The 120 kHz represented a slightly larger area and therefore had the possibility of detecting seabed substrate properties that the 38 kHz did not. However, this issue is not considered significant and will not affect the data in anyway that can be demonstrated as a footprint is at best only within 6 m of its true position. The one-degree difference between incidence angles used to define near-nadir backscatter of both frequencies is not considered significant as these angles were chosen to standardize backscatter to its peak amplitude response. The second issue was that there was a slight offset (0.5 seconds in 2002 and 0.25 seconds in 2003) in the projection of sound waves between the two frequencies in order to reduce noise (section 3.2). Therefore, the footprints of each pair of dual-frequency samples did not exactly spatially overlap, however much of the analysis (e.g. quantile classifications, Wilcoxon tests, regressions, correlations, k-means, discriminant function analysis) required using datasets in which one sample had both a 38 kHz and 120 kHz value. To achieve this overlap, the geographic coordinates from each frequency were averaged to create spatially 'balanced' datasets. The new sample location representing data from the two frequencies was well inside the original footprints of both frequencies, with a shift ranging from 1 m – 5 m along-track compared to footprint sizes of 14 m – 20 m along track. Again, this shift should not significantly affect any results as there is a \pm 3m positional accuracy for the original acoustic footprints.

Expected measurement errors within and between surveys, resulting from equipment, sampling, and environmental differences were controlled for as much as possible. The parameters of the acoustic system, including frequency, pulse length,

transmit power, and sampling rate were held constant. Vessel speed was also held constant throughout the surveys. A CTD cast done before each survey provided information to measure sound speed and the absorption coefficient. Post-processing of the data compensated for TVG, absorption, changes in depth, incidence angle, and beam pattern (see section 3.3). Ship motion was not considered; however, all sampling was done under very calm sea states (Beaufort Scale 0) and was not considered to vary significantly between areas or years. A formal quantification of the errors would have required more ground-truthing, which was not possible given the ship time available and the more general survey design. Finally, although rescaling backscatter values for differences in acoustic systems and years was necessary, the data are still comparable between years and systems because the acoustic systems were calibrated to specific standards and the data were standardized for radiometric and geometric biases.

The final consideration is the use of near-nadir backscatter only, which was based on theoretical explanations and previous work on the Scotian Shelf (Courtney et al. (2005). Near-nadir backscatter results from several processes: acoustic impedance contrast, surface roughness at comparable scales to the wavelength, and volume scattering caused by inhomogeneities in the substrate volume. Models show that near-nadir backscatter has complex relationships between these factors and delineating seabed types may be limited by using the near-nadir response only (APL94, 1994). Off-nadir response could be explored in the future to gain additional information for delineating seabed types. Even though Courtney et al. (2005) concluded that the peak amplitude response was the driving factor in discriminating between sand and gravel substrate for this study area, they do acknowledge that this may not apply to areas of different substrate types.

6. Conclusions

This research has compared near-nadir acoustic backscatter from single beam echosounders using two acoustic frequencies on the Scotian Shelf, Canada, and evaluated acoustic seabed classification (ASC) results using single frequency and dual-frequency data (38 kHz and 120 kHz). Previous studies indicated that the shape and intensity of seabed backscatter vary as a function of physical seabed properties (grain size, porosity, and roughness) and the acoustic frequency (amplitude and wavelength). Therefore, it was hypothesized in this research that combining near-nadir backscatter responses from two acoustic frequencies would provide additional information that will help to improve our understanding of seabed types and therefore improve ASC. Research questions, objectives, and methods were designed to address this hypothesis and investigate its validity.

The goal of this research was to measure and interpret variations in acoustic backscatter values as a function of frequency and seabed substrate characteristics at different spatial scales on Western Bank and to determine if using dual-frequency backscatter could improve ASC. This was approached by statistically analyzing and classifying near-nadir backscatter values for each frequency independently, and in combination, and comparing results to interpreted surficial geology maps. Specific frequency differences between backscatter values were highlighted, and both univariate and multivariate classification techniques were used to perform ASC. The methods used were very much shaped by the research objectives outlined in the introduction, and in turn, these research objectives were shaped by initial research questions, which are restated below:

1. What are the univariate (single frequency) statistical characteristics of near-nadir acoustic backscatter data from each frequency independently and do they differ?
2. How do univariate classifications of near-nadir acoustic backscatter values from each frequency relate to seabed substrate types?
3. Where do the major frequency differences occur and what are the spatial scales at which these frequency differences take place?
4. What are the multivariate statistical characteristics of near-nadir acoustic backscatter data from both frequencies?
5. How do multivariate classifications of dual-frequency near-nadir acoustic backscatter values relate to seabed substrate types?
6. Does dual-frequency near-nadir acoustic backscatter improve the ability to classify seabed substrate types compared to a single acoustic frequency?

The methods and results used to address these research questions were examined at three different levels of observation: 1) results for all data from both study area, 2) results for data from all combined ISUs with dominant sand substrate and all combined ISUs with dominant gravel substrate, and 3) results for data from each ISU independently.

Question 1 was answered using univariate (i.e. single frequency) descriptive statistics. The main conclusions were that 120 kHz generally produced higher near-nadir backscatter values than 38 kHz, gravel seabeds generally produced higher backscatter values than sand seabeds, and 120 kHz provided greater differences in backscatter values between seabeds of different substrate type and therefore was likely to provide better discrimination between surficial seabed substrate types. These results were expected based on existing literature (see discussion in chapter 5). In addition, seabeds with mixed sediments or surficial roughness at comparable scales to the acoustic footprint such as sand to sand gravel (SG), sand with boulders (SB), gravel ripples – short wavelength (GRS), and gravel ripple incised (GRI) generally had higher variance

values and were therefore less likely to conform to the general trend of higher near-nadir backscatter values from gravel substrate. Frequency distribution plots of the quantile classification results used to address question 2 also reflected these main results.

Research question 2 inquired about the ability of a single frequency to differentiate seabed types. Quantile classifications conveyed that it was difficult to classify the ISUs by simply using the near-nadir component of backscatter from a single frequency. These results were consistent with previous single frequency classification results (Courtney et al., 2005) that could only differentiate, at best, between the two dominant substrate types on the Western Bank (i.e. sand and gravel). There were several reasons why the near-nadir backscatter could not differentiate between the seven or eight different ISUs in both study areas. The ISUs were classified using much higher resolution acoustic data than used for this project, and the footprint of the SBES could not detect all the varying degrees of roughness and topographic features of the ISUs. This provided reasoning for the use of additional morphology variables in the multivariate classifications.

Research question 3, concerning the presence and scale of frequency differences, was answered using both classified (i.e. normalized) and continuous near-nadir backscatter values. Frequency differences based on classified backscatter values revealed that the 38 kHz frequency produced higher classified backscatter from sand seabeds and the 120 kHz frequency produced higher classified backscatter from gravel seabeds. This could be a result of the 38 kHz frequency penetrating deeper into the sand seabeds and detecting underlying gravel, which would have produced higher classified backscatter values. However, quantitative measurements of the amount of surface and volume scattering would have provided better support for this assumption. Frequency differences of continuous backscatter values were analyzed over the scale of tens to

hundreds of meters, and fewer frequency differences occurred over larger scales. Therefore, the scale at which frequencies are compared is very important. It was difficult, however, to associate any frequency differences with specific seabed classes or ISUs.

Question 4 was answered using a number of multivariate statistical analysis using near-nadir backscatter values from both frequencies. Correlation values revealed that there was a positive relationship between the two frequencies, but it was not particularly strong. The explained variation values from the linear regression models were low. This led to the conclusion that the two frequencies did not completely share a one to one relationship, which potentially adds information when using two acoustic frequencies. In addition, there doesn't seem to be specific acoustic signatures of certain ISUs when analyzing backscatter values from both frequencies using linear regression analysis.

Research questions 5 and 6 were specifically answered by using multivariate classification. The multivariate classification results did slightly better than single frequency results, but this is probably not significant. Discriminant function analysis provided classification accuracy values for each ISU. The overall classification accuracies of supervised classifications using dual-frequency near-nadir backscatter and morphology variables were not particularly high. Supervised classification results could not do better than 50% accuracy. Nevertheless, there were certain ISUs that had much higher classification accuracies than others. Gravel ripples – short wavelength (GRS), gravel ripples – long wavelength (GRL), gravel lag (GL), and sand with megaripples in the non-preferred study area, and gravel ripples incised (GRI), gravel lag (GL), and Sand in the preferred study area all had classification accuracies higher than 50%. Univariate classification using one frequency could at best discriminate between sand and gravel seabeds based on high and low backscatter values, whereas multivariate classification

results had three or more ISUs with classification accuracies greater than 50%. Multivariate classifications using one frequency indicated that the 120 kHz generally provided better discrimination between ISUs. In addition, multivariate classification results had higher classification accuracies when using two frequencies as opposed to one. Therefore, it can be concluded that dual-frequency classification of seabed types performs better than single frequency classification.

The main results from the thesis are summarized here:

1. 120 kHz near-nadir backscatter was generally higher than 38 kHz near-nadir backscatter
2. Gravel substrate produced higher near-nadir backscatter than sand substrate
3. Certain ISUs with mixed sediments or rougher surfaces generally tended to produce variable results, not consistent with typical near-nadir backscatter trends of dominant sand and gravel substrate ISUs, and lower classification accuracies
4. 120 kHz provided better discrimination between substrate types
5. Frequency differences indicated that 38 kHz was higher over sand seabeds and 120 kHz was higher over gravel seabeds in terms of classified backscatter values
6. Frequency differences of both classified and continuous backscatter values occurred the most over the sand seabeds in the non-preferred study area
7. The scale at which frequency differences were observed occurred across hundreds of meters and dissipated at the scale of thousands of meters

8. Single frequency classifications could at best discriminate between sand and gravel substrate based on high and low backscatter values
9. The preferred study area had the lowest supervised classification accuracy, and therefore proved difficult to discriminate between some seabed substrate types. This could be due to the higher variability in the number and distribution of ISUs in this study area.
10. Dual-frequency classification improved slightly upon single-frequency classification of surficial seabed types

Conclusions 1 to 3 were expected results as these facts were well known, while conclusion 4 confirms results from previous studies of multi-frequency backscatter. With respect to conclusion 3, one must be cautious when generalizing the seabed into sand and gravel seabed types as ISUs with heterogeneous sediments or rougher surfaces can create large variability within a generalized class. Conclusion 8 reiterates the fact that a more complex environment like the seabed in the preferred study area proves difficult to perform acoustic seabed classification well. Finally, the fact that dual-frequency classification improved upon single frequency classification may not be significant as the improvement was not large.

Part of the hypothesis stated that backscatter from multiple frequencies should provide additional information on the seabed. This was validated through a number of different results. The fact that the ranking of mean near-nadir backscatter values of the ISUs from highest to lowest varied between frequencies indicate that backscatter from the two frequencies vary differently and therefore may be detecting different physical properties. The analysis of frequency differences and the low coefficient of determination (r^2) values from the linear regressions also convey that there are differences in near-nadir backscatter from the two frequencies that could provide additional information on

the seabed. The remaining part of the hypothesis stated that dual-frequency backscatter should improve ASC. It was demonstrated that dual-frequency backscatter improved slightly upon single-frequency ASC results through discriminant function analysis results. Hence, it was concluded that the physical properties of the seabed were better represented through two different frequencies, although further work would be necessary to better understand the processes observed.

There certainly are some important issues to address here in this research and future research paths that can be taken. As acoustic backscatter data are rather noisy, it may have been useful to use backscatter data that were smoothed (for example using a running average) in order to reduce variability. Statistical and classification results from this may have been easier to interpret. Regardless, one must be aware of the uncertainties and expected errors in acoustic backscatter measurements. It would have also been useful to have measures of the amount of surface and volume scattering from both frequencies in order to make inferences about some of the observed patterns (e.g. higher classified backscatter from the 38 kHz frequency observed over sand seabeds). Another important issue pertains to the interpreted sediment units (ISUs) of the interpreted surficial geology maps and their use as a ground-truthing tool (these issues are discussed in chapter 5). Perhaps more direct, accurate, and detailed ground-truthing of seabed substrate providing well characterized reference sites of known geacoustic properties would have been useful in detecting frequency variations. Sediment grabs and particle size analysis of specific areas could have been correlated with backscatter values of both frequencies in order to characterize backscatter in relation to grain size and frequency; however, these data were not available. Finally, the study areas mostly had substrate ranging from sand to gravel. Having studied an area having a broader range of substrate types may have provided further insight into the usefulness of multi-

frequency backscatter. The relationship between substrate type and frequency is obvious and consistent differences in frequencies may have been easily identified in such an environment.

Concerning the classifications, the near-nadir backscatter data and morphology layers used were not expected to be able to detect all the fine details of the roughness and grain size distributions of the ISUs. Hence, the classification accuracies were dependent upon the input that was given to the classification algorithm, and this must be considered when interpreting the results. The morphology layers (slope and curvature) used for input into multivariate classification were chosen because they were able to measure simple roughness of the seabed. However, there are a myriad of other morphology variables that may have been used for this purpose (see Wilson et al., 2007 for example). Fuzzy classification could have been used as an alternative to the classification schemes used here (Lucieer and Lucieer, 2009). This type of classification would provide useful information on the probability values of having the correct classification for each observation.

The project used 38 kHz and 120 kHz frequencies. While the use of lower (e.g. 12 kHz) or higher (e.g. 200 kHz) frequencies could have been interesting to analyse, 38 kHz and 120 kHz were the only two frequencies collected by Fisheries and Oceans Canada in these study areas. Off-nadir backscatter could have been a useful measure to use here as well, as it could have provided more information on the seabed in addition to the near-nadir backscatter based on certain model results (APL94, 1994). Finally, this study was an empirical comparison between two frequencies without having done any comparative modelling work. Hence, it is not known if greater backscatter differences between frequencies would occur as a function of incidence angle. It was expected that the dominant frequency dependent response would be found at near nadir (Courtney et

al., 2005). A comparative modelling study matching empirical and model analyses to test classification results could be considered as a future research path.

The results from this research are significant for a number of reasons. Despite the increasing number of acoustic systems that can acquire data using different frequencies, there have been a limited number of research projects that attempt to classify seabed environments by combining more than one acoustic frequency. This study has also used transparent statistical analysis instead of black-box commercial technology. Therefore, the results from this research should be useful to researchers who would like to know how acoustic seabed classification performs under multi-frequency analysis. However, it is important to point out that these results have shown that there are difficulties in discriminating seabed types using the near-nadir backscatter component only, and this is reflected in model results (APL94, 1994). The research did highlight significant near-nadir backscatter differences as a function of acoustic frequency at scales of hundreds to thousands of meters. In addition, the two frequencies used in this project are commonly used worldwide for seabed classification projects, fisheries and biological assessments, and habitat mapping efforts. Hence, researchers and agencies (such as Fisheries and Oceans Canada) in these fields could use these results to better understand how near-nadir 38 kHz and 120 kHz backscatter can be characterized and how well it performs in a classification scheme.

References

- Amos, C.L., and King, E.L. (1984). Bedforms of the Canadian eastern seaboard: a comparison with global occurrences. *Marine Geology*, 57, 167-206.
- Anderson, J.T. (2007). Acoustic seabed classification of marine physical and biological landscapes. In ICES Cooperative Research Report No. 286, *Acoustic Seabed Classification of Marine Physical and Biological Landscapes*, (pp. 1-6).
- Anderson, J.T., Gregory, R., and Collins, W. (2002). Acoustic classification of marine habitats in coastal Newfoundland. *ICES Journal of Marine Science*, 59(1), 156-167.
- Anderson, J.T., Simon, E., Gordon, D., and Hurley, P. (2005). Linking fisheries to benthic habitats at multiple scales: Eastern Scotian Shelf Haddock. *American Fisheries Society Symposium*, 41, 251-264.
- Anderson, J.T., Courtney, R.C., and Fader, G.B.J. (2007). Acoustic surrogates for demersal fish habitats on the Scotian Shelf: Juvenile Haddock and Atlantic Cod. In Canadian Technical Report of Report of Fisheries and Aquatic Sciences No. 2770, *Report of the Spatial Utilization of Benthic Habitats by Demersal Fish on the Scotian Shelf Synthesis Meeting*, (pp. 32-33).
- Anderson, J.T., Van Holliday, D.V., Kloser, R., Reid, D.G., and Simard, Y. (2008). Acoustic seabed classification: current practice and future directions. *ICES Journal of Marine Science*, 65(6), 1004-1011.
- APL94-Applied Physics Laboratory. (1994). High frequency ocean environmental acoustic model handbook APL-UW TR 9407. *Applied Physics Laboratory Technical Report* (University of Washington).
- Brown, C.J., and Collier, C.S. (2008). Mapping benthic habitat in regions of gradational substrate: an automated approach using geophysical, geological, and biological relationships. *Estuarine, Coastal and Shelf Science*, 78, 203-214.
- Buchanan, J., Gilbert, R., Wirgin, A., and Yongzhi, X. (2004). *Marine acoustics: direct and inverse problems*. Philadelphia, PA: Society for Industrial and Applied Mathematics.
- Chakraborty, B., Mahale, V., Navekar, G., Rao, B.R., Prabhudesai, R.G., Ingole, B., and Janakiram, G. (2007). Acoustic characterization of seafloor habitats on the western continental shelf of India. *ICES Journal of Marine Science*, 64(3), 551-558.
- Clay, C.S., and Leong, W.K. (1974). Acoustic estimates of the topography and roughness spectrum of the sea floor southwest of the Iberian Peninsula. In L. Hampton (Ed.), *Physics of sound in marine sediments* (pp. 373-446). Plenum Press, New York, NY.

- Cogan, C.B., Todd, B.J., Lawton, P., and Thomas, T.N. (2009). The role of marine habitat mapping in ecosystem-based management. *ICES Journal of Marine Science*, 66, 2033-2042.
- Collier, J.S., and Brown C.J. (2005). Correlation of sidescan backscatter with grain size distribution of surficial seabed sediments. *Marine Geology*, 214, 431-449.
- Collins, W.T., and Rhynas, K.P. (1998). Acoustic seabed classification using echo sounders: operational considerations and strategies. In *Proceedings of the Canadian Hydrographic Conference*, Victoria, BC, Canada, 10-12 March 1998.
- Courtney, R., Anderson, J.T., Lang, C., and Fader, G.B.J. (2005). Comparative seabed classification using sidescan and normal incidence sonar data at selected study sites on the Scotian shelf, Canada. In *Proceedings of the International Conference "Underwater Acoustic Measurements: Technologies and Results"*, Crete, Greece, 28 June - 1 July 2005.
- DOSITS, Discovery of Sound in the Seas. (2009). Retrieved from <http://www.dosits.org/>
- Durand, S., Legendre, P., and Juniper, S.K. (2006). Sonar backscatter differentiation of dominant macrohabitat types in a hydrothermal vent field. *Ecological Applications*, 16(4), 1421-1435.
- Echoview (Myriax). (2010). Echoview Help File for Echoview release 4.90.44.
- ESRI ArcGIS. (2008). Desktop Help for ArcGIS release 9.3.
- Fader, G.B.J. (2007). Seabed sediment distributions, morphology, dynamics and features of detailed study areas on Emerald, Western and Sable Island Banks, Outer Scotian Shelf. In Canadian Technical Report of Report of Fisheries and Aquatic Sciences No. 2770, *Report of the Spatial Utilization of Benthic Habitats by Demersal Fish on the Scotian Shelf Synthesis Meeting*, (pp. 29-31).
- Ferrini, V.L., and Flood, R.D. (2006). The effects of fine-scale surface roughness and grain size on 300 kHz multibeam backscatter intensity in sandy marine sedimentary environments. *Marine Geology*, 228, 153-172.
- Fossa, J.H., Lindberg, B., Christensen, O., Lundelv, T., Svellingen, I., Mortensen, P.B., and Alvsvag, J. (2005). Mapping of *Lophelia* reefs in Norway: experiences and survey methods. In A. Freiwald and Roberts J.M. (Eds), *Cold-water Corals and Ecosystems* (pp. 337-370). Berlin Heidelberg, Germany: Springer-Verlag.
- Foster-Smith, R.L., and Sothoran, I.S. (2003). Mapping marine benthic biotopes using acoustic ground discrimination systems. *International Journal of Remote Sensing*, 42(13), 2761-2784.
- Foot, K. G., Knudsen, H. P., Vestnes, G., MacLennan, D. N., and Simmonds, E. J. (1987). Calibration of acoustic instruments for fish density estimation: a practical guide. ICES Cooperative Research Report No. 144, 69 pp.

Friel, C.M. (2009). Sam Houston State University, Criminal Justice Center. *Statistical Digests Homepage*, CJ 742 Advanced Statistics I, Discriminant Analysis. Retrieved October, 28, 2009, from http://www.shsu.edu/~icc_cml/cj_742.html

Galloway, J.L., and Collins, W.T. (1996). Dual frequency acoustic classification of seafloor habitat using the QTC VIEW. In *Proceedings of the IEEE Oceanic Engineering Society Conference "Oceans '98"*, 3, Nice, France, 28 September – 1 June 1998 (pp. 1296 – 1300). USA: Institute of Electrical and Electronics Engineers (IEEE) Inc & the Oceanic Engineering Society (OES).

Garson, G.D. (2008). North Carolina State University, College of Humanities and Social Sciences. Quantitative Research in Public Administration Course. Retrieved September 5, 2009, from <http://faculty.chass.ncsu.edu/garson/PA765/weekly.htm>

Gordon, D.C., Kenchington, E.L.R., Gilkinson, K.D., McKeown, D.L., Steeves, G., Chin-Yee, M., Vass, W.P., Bentham, K., and Boudreau, P.R. (2000). Canadian imaging and sampling technology for studying marine benthic habitat and biological communities. ICES CM 2000/T:07, 11p.

Gregory, R.S., and Anderson, J.T. (1997). Substrate selection and use of protective cover by juvenile Atlantic cod *Gadus morhua* in inshore waters of Newfoundland. *Marine Ecology Progress Series*, 146, 9-20.

Hamilton, E.L., and Bachman, R.T. (1982). Sound velocity and related properties of marine sediments. *Journal of the Acoustical Society of America*, 72(5), 1891-1904.

Holliday, D.V. (2007). Theory of sound-scattering from the seabed. In *ICES Cooperative Research Report No. 286, Acoustic Seabed Classification of Marine Physical and Biological Landscapes*, (pp. 7-28).

Jackson, D.R., Baird, A.M., Crisp, J.J., and Thomson, P.A.G. (1986). High frequency bottom backscatter measurements in shallow water. *Journal of the Acoustical Society of America*, 80(4), 1188-1199.

Jackson, D.R., and Briggs, K. (1992). High-frequency bottom backscattering: roughness versus sediment volume scattering. *Journal of the Acoustical Society of America*, 92(2), 962-977.

Jackson, D.R., Briggs, K.B., Williams, K.L., and Richardson, M.D. (1996). Tests of models for high-frequency seafloor backscatter. *IEEE Journal of Oceanic Engineering*, 21(4), 458-470.

Jackson, D.R., and Richardson, M.D. (2007). *High-Frequency Seafloor Acoustics*. New York, New York: Springer Science.

Jensen, J.R. (2005). *Introductory digital image processing: a remote sensing perspective* (3rd edition). Upper Saddle River, NJ: Pearson Prentice Hall.

- Kenny, A., Cato, I., Desprez, M., Fader, G., Schuttenhelm, R., and Side, J. (2003). An overview of seabed-mapping technologies in the context of marine habitat classification. *ICES Journal of Marine Science*, 60(2), 411-418.
- Kloser, R. (2007). Seabed backscatter, data collection, and quality overview. in ICES Cooperative Research Report No. 286, *Acoustic Seabed Classification of Marine Physical and Biological Landscapes*, (pp. 45-60).
- Kloser, J., Bax, N.J., Ryan, T., Williams, A., and Barker, B.A. (2001). Remote sensing of seabed types in the Australian south east fishery: development and application of normal incident acoustic techniques and associated 'ground-truthing'. *CSIRO Marine & Freshwater Research*, 52, 475-489.
- Kloser, R., Keith, G., Ryan, T., Williams, A., and Penrose, J. (2002). Seabed biotope characterisation in deep water – initial evaluation of single and multi-beam acoustics. In A. Stepienowski, (Eds). *Proceedings of the sixth European conference on underwater acoustics* (pp. 81-88). Gdansk, Poland.
- Kornellussen, R., and Ona, E. (2002). An operational system for processing and visualizing multi-frequency acoustic data. *ICES Journal of Marine Science*, 59(2), 293-313.
- Kostylev, V., Todd, B., Fader, G., Courtney, R., Cameron, G., and Pickill, R. (2001). Benthic habitat mapping on the Scotian Shelf based on multibeam bathymetry, surficial geology and sea floor photographs. *Marine Ecology Progress Series*, 219, 121-137.
- Leech, N., Barnett, K., and Morgan, G. (2005). *SPSS for intermediate statistics: use and interpretation*, (2nd edition). Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.
- Legendre, P. (2001). Program K-means user guide. Département de sciences biologiques, Université de Montréal. 11 pages.
- Legendre, P., and Legendre, L. (1998). *Numerical Ecology* (2nd Edition). Amsterdam, The Netherlands: Elsevier Science.
- Legendre, P., Ellingsen, K., Bjørnø, E., and Casgrain P. (2002). Acoustic seabed classification: improved statistical method. *Canadian Journal of Fisheries and Aquatic Sciences*, 59(7), 1065-1089.
- Lucieer, V., and Lucieer, A. (2009). Fuzzy clustering for seafloor classification. *Marine Geology*, 264, 230-241.
- Lurton, X. (2002). *An introduction to underwater acoustics: principles and applications*. Chichester, UK: Praxis Publishing Ltd.
- Manik, H.M., Furusawa, M., and Amakasu, K. (2006). Measurement of sea bottom surface backscattering strength by quantitative echo sounder. *Fisheries Science*, 72(3), 503-512.

McGrew, J.C., and Monroe, C.B. (2000). *An introduction to statistical problem solving in geography* (2nd edition). United States: McGraw-Hill Higher Education.

McKinney, C.M., and Anderson, C.D. (1964). Measurements of backscattering of sound from the ocean bottom. *The Journal of the Acoustical Society of America*, 36(1), 158-163.

Nafe, J. E., and Drake, C. L. (1964). Physical properties of marine sediments. In M. N. Hill (Ed.), *The sea*, Vol. 3 (pp. 794-815). New York, NY: John Wiley and Sons.

NOAA - National Oceanic and Atmospheric Organization. (2009). Coastal Services Center: benthic habitat mapping website. Retrieved April 8, 2009 from <http://www.csc.noaa.gov/benthic/>

Ollerhead, N., and Anderson, J.T. (2007). Habitat suitability criteria for juvenile Haddock and Atlantic Cod on the Scotian Shelf. In Canadian Technical Report of Report of Fisheries and Aquatic Sciences No. 2770, *Report of the Spatial Utilization of Benthic Habitats by Demersal Fish on the Scotian Shelf Synthesis Meeting*, (pp. 49-52).

Parrott, D.R., Dodds, D.J., King, L.H., and Simpkin, P.G. (1980). Measurement and evaluation of the acoustic reflectivity of the seafloor. *Canadian Journal of Earth Science*, 17(6), 722-737.

Pike, J. (1998). Underwater Acoustics. Retrieved from <http://www.fas.org/man/dod-101/sys/ship/acoustics.htm>

Phon, B.S., and Chriaman, N.R. (2006). Order from noise: toward a social theory of geographic information. *Annals of the Association of American Geographers*, 96(3), 508-523.

Riegl, B.M., and Purkis, S.J. (2005). Detection of shallow subtidal corals from IKONOS satellite and QTC View (50, 200 kHz) single-beam sonar data (Arabian Gulf; Dubai, UAE). *Remote Sensing of Environment*, 95, 96-114.

Roberts, J.M., Brown, C.J., Long, D., and Bates, C.R. (2005). Acoustic mapping using a multibeam echosounder reveals cold-water coral reefs and surrounding habitats. *Coral Reefs*, 24, 654-669.

Shaw, P.J. (2005). *Multivariate statistics for the environmental sciences*. New York, NY: Oxford University Press.

Simard, Y., and Stepnowski, A. (2007). Classification methods and criteria. In ICES Cooperative Research Report No. 286, *Acoustic Seabed Classification of Marine Physical and Biological Landscapes*, (pp. 61-72).

Simmonds, J., and MacLennan, D. (2005). *Fisheries Acoustics: Theory and Practice* (2nd Edition). Cornwall, Great Britain: Blackwell Science Ltd.

SPSS Inc. (2007). Online help system for SPSS release 16.0.1.

Sternlicht, D.D., and de Moustier, C.P. (2003). Time-dependent seafloor acoustic backscatter (10-100 kHz). *Journal of the Acoustical Society of America*, 114(5), 2709-2725.

Tabachnick, B.G., and Fidell, L.S. (2001). *Using multivariate statistics* (4th edition). Needham Heights, MA: Allyn & Bacon.

United Nations Convention on the Law of the Sea (UNCLOS). (December 10th, 1982). Retrieved May 18, 2008 from <http://www.un.org/Depts/los/index.htm>

Appendix A: Interpreted Sediments Units (ISUs)

Interpreted Sediment Units (ISUs) for three fishing banks (Emerald, Sable Island, and Western) on the Scotian Shelf (Source: Fader, 2007). ISUs present on Western Bank are highlighted in both the preferred and non-preferred study area columns.

Interpreted Sediment Unit (ISU)	ISU #	Preferred Study Area	Non-preferred Study Area
GRAVEL - Gravel	1		
GS - Gravel to gravely sand	2		
GR - Gravel ripples	3		
GRS - Gravel ripples, short wavelength	4	X	X
GRI - Gravel ripples, long wavelength	5	X	X
GRI - Gravel ripples, incised	6	X	X
GL - Gravel lag	7	X	X
GH - Gravel, hummocky	8		X
GSP - Gravel with small sand patches	9		
GSRT - Gravel with sand ribbon troughs	10		
SAND - Sand	11	X	X
SM - Sand with megaripples	12	X	X
SG - Sand to sandy gravel	13	X	
SB - Sand with scattered boulders	14		X

Appendix B: Mann-Whitney U Test Results

Wilcoxon W and Mann-Whitney U Analysis for the Non-preferred Study Area (120 kHz). Non-significant differences highlighted in bold font. The values in the ISU # column refer to the ISU # from the table in Appendix A.

ISU #	N	Mean Rank	Sum of Ranks	Mann-Whitney U	Wilcoxon W	Z	Significance Level
4	66	35.20	2323				
5	14	85.50	817	112.00	2323.00	-4.432	0.000
4	56	70.86	4878	2467.00	4878.00	-0.680	0.496
6	30	75.66	8953				
4	56	45.82	3628	817.00	3628.00	-9.526	0.000
7	140	133.68	18293				
4	56	96.17	6347				
8	229	162.94	37313	4136.00	6347.00	-5.603	0.000
4	56	256.06	17032				
11	547	312.80	171158	14621.00	17032.00	-2.377	0.017
4	66	234.81	15604	13293.00	15504.00	-4.762	0.000
12	627	256.80	224967				
4	66	48.33	3190	878.00	3190.00	-2.223	0.026
14	40	62.02	2481				
9	14	82.03	1165	64	3304	-5.267	0.000
9	80	41.30	3304				
9	14	58.00	770	965	770	-1.980	0.048
7	140	78.75	11168				
5	14	183.64	2281	1820	27305	-2.263	0.022
9	229	119.45	27355				
5	14	532.37	3449	314	190192	-5.869	0.000
15	547	274.57	150150				
5	14	556.30	7817	1386	187944	-4.849	0.000
12	627	318.70	197944				
5	14	38.07	547	118	936	-3.198	0.001
14	40	23.48	936				
6	80	45.71	3657	417	3657	-11.412	0.000
7	140	147.52	20653				
6	80	94.22	7536	4256	7536	-7.067	0.000
8	229	176.23	40367				
6	80	278.21	22241	19801	22241	-1.962	0.057
11	547	318.26	174637				
6	80	231.81	18545	15305	18545	-5.682	0.000
12	627	368.89	231753				
6	80	55.15	4432	1172	4432	-2.383	0.017
16	40	71.20	2848				
7	140	267.27	36918	5912	32247	-10.177	0.000
8	229	148.82	32247				
7	140	400.88	84121	2329	152207	-17.162	0.000
11	547	278.26	152207				
7	140	497.44	82041	5609	202487	-18.151	0.000
12	627	322.96	202487				
7	140	106.38	14793	7137	1537	-7.167	0.000
14	40	36.42	1537				
8	229	527.93	120896	36702	180580	-11.211	0.000
11	547	330.13	180580				
8	229	531.34	121676	48242	245120	-7.354	0.000
12	627	390.94	245120				
8	229	159.57	31962	3533	4353	-2.306	0.021

14	40	136.87	4353				
11	547	487.56	206037	116879	268897	-6.433	0.000
12	627	674.89	423026				
11	547	290.11	158689	8811	158689	-2.056	0.048
14	40	367.22	13889				
12	627	334.08	209470	12486	13308	-1.044	0.965
14	40	332.70	13308				

Wilcoxon W and Mann-Whitney U Analysis for the Non-preferred Study Area (38 kHz). Non-significant differences highlighted in bold font. The values in the ISU # column refer to the ISU # from the table in Appendix A.

ISU #	N	Mean Rank	Sum of Ranks	Mann-Whitney U	Wilcoxon W	Z	Significance Level
4	61	32.79	1999	107	1999	-4.574	0.000
5	15	67.87	1018				
6	61	94.26	5757	1406	3297	-4.041	0.000
6	77	61.74	4754				
6	61	42.77	2609	718	2609	-5.919	0.000
7	123	117.81	14471				
4	61	86.43	4806	3015	4806	-6.281	0.000
8	209	151.57	31679				
4	61	189.93	11586	9695	11586	-5.342	0.000
11	548	336.21	172335				
4	61	146.10	8912	7021	8912	-6.142	0.000
12	624	262.25	226043				
6	61	42.87	2607	730	2621	-2.920	0.003
14	37	66.27	2335				
5	15	71.73	1076	199	3202	-4.091	0.000
6	77	61.86	3202				
5	15	71.27	1069	896	8522	-0.181	0.856
7	123	66.26	8122				
5	15	181.87	2609	827	22772	-0.054	0.002
9	209	186.96	32772				
5	15	475.40	7131	1164	148949	-4.726	0.000
11	548	275.14	148949				
5	15	489.20	7338	2442	197442	-3.158	0.002
12	624	336.41	197442				
5	15	36.53	548	127	830	-0.040	0.002
14	37	22.43	830				
6	77	58.73	4291	1288	4291	-6.656	0.000
7	123	128.53	15809				
6	77	136.25	8951	5948	8951	-3.383	0.001
8	209	153.54	32090				
6	77	378.68	29528	25430	168215	-0.374	0.709
11	548	370.49	168215				
6	77	236.77	18385	15382	18385	-5.155	0.000
12	624	364.85	227686				
6	77	56.71	4367	1364	4367	-0.366	0.714
14	37	59.14	2188				
7	123	220.37	27126	3227	28172	-7.846	0.000
8	209	134.79	28172				
7	123	554.87	68249	6432	158187	-14.021	0.000
11	548	284.77	155187				
7	123	545.99	67127.5	17270	212270.5	-6.649	0.000
12	624	340.18	212170.5				
7	123	81.38	10015	962	1866	-6.376	0.000
14	37	69.30	1866				
8	209	485.79	97340	38500	197285	-5.893	0.000

11	545	343.64	187285				
8	298	433.00	84228	62283	84228	-0.972	0.331
12	624	421.69	263133				
9	298	126.36	28415	3268	3871	-1.5	0.134
14	37	187.32	3971				
10	545	444.74	242382	93587	242382	-13.275	0.000
12	624	797.90	441483				
11	545	290.27	158088	9303	158088	-0.788	0.431
14	37	312.87	11668				
12	624	335.15	209132	8956	9659	-2.293	0.022
14	37	281.85	9659				

Wilcoxon *W* and Mann-Whitney *U* Analysis for the Preferred Study Area (120 kHz). Non-significant differences highlighted in bold font. The values in the ISU # column refer to the ISU # from the table in Appendix A.

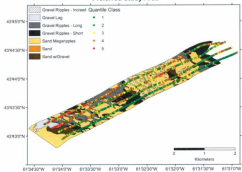
ISU #	N	Mean Rank	Sum of Ranks	Mann-Whitney U	Wilcoxon W	Z	Significance Level
4	275	236.90	56037	18947	56897	-11.085	0.000
5	297	380.27	106981				
4	275	180.21	52232	7698	10908	-6.087	0.000
5	67	135.73	10698				
4	275	141.27	35795	845	38795	-8.091	0.000
7	38	270.03	8727				
4	275	138.45	38073	123	38073	-3.146	0.002
11	5	283.40	1267				
4	275	554.55	152402				
12	737	488.57	360076	58123	360076	-3.195	0.001
4	275	175.17	48172	10222	48172	-3.676	0.000
13	99	221.35	21953				
5	297	218.80	64945	3028	6268	-19.232	0.000
6	63	78.35	4268				
5	297	158.11	46948	2735	46956	-4.841	0.000
7	38	240.38	8853				
5	297	180.88	44870	557	44870	-8.98	0.338
11	5	188.60	943				
5	297	743.19	220728	42476	214389	-15.427	0.000
12	737	435.55	314369				
5	297	211.42	62782	10864	15874	-3.891	0.000
13	99	159.74	15874				
6	63	41.81	2345	105	3345	-7.967	0.000
7	38	95.56	3441				
6	63	40.90	2547	7	3247	-3.605	0.000
11	5	81.80	408				
6	63	360.19	28075	24775	28075	-2.347	0.019
12	737	415.38	308138				
6	63	85.32	5218	1970	5218	-5.774	0.000
13	99	130.13	12990				
7	38	22.86	796	52	67	-1.574	0.120
11	5	53.40	87				
7	38	794.22	25244	1954	273967	-8.647	0.000
12	737	371.68	273967				
7	38	38.25	3537	693	5643	-5.419	0.000
13	99	57.90	5643				
11	5	680	3438	276	272223	-3.292	0.001
12	737	369.37	272223				
11	5	75.80	393	117	5067	-1.983	0.047
13	99	51.18	5067				
12	737	400.99	295463	23510	295463	-6.750	0.000
13	99	549.53	54433				

Wilcoxon W and Mann-Whitney U Analysis for the Preferred Study Area (38 kHz). Non-significant differences highlighted in bold font. The values in the ISU # column refer to the ISU # from the table in Appendix A.

ISU #	N	Mean Rank	Sum of Ranks	Mann-Whitney U	Wilcoxon W	Z	Significance Level
4	276	276.30	59130	21474	58700	-9.208	0.000
5	283	342.12	96820				
4	276	170.07	46940	8714	46940	-3.722	0.000
8	88	218.17	18763				
4	276	142.34	39259	1033	39259	-6.905	0.000
7	31	258.68	8019				
4	276	139.42	38480	254	38480	-2.421	0.015
11	5	228.20	1141				
4	276	479.82	118670	77644	118670	-5.777	0.000
12	736	539.01	396706				
4	276	187.83	46320	8094	46320	-5.906	0.000
13	98	242.91	23805				
8	283	184.76	52116	3408	13149	-3.187	0.001
8	88	152.80	13149				
5	283	149.65	42352	2186	42352	-4.627	0.000
7	31	229.13	7103				
5	283	143.82	40730	814	40730	-1.048	0.295
11	5	183.20	915				
5	283	503.90	117094	77570	348786	-6.316	0.000
12	736	473.69	348786				
5	283	184.90	52157	12763	17614	-1.175	0.240
13	98	179.73	17614				
8	88	48.84	4200	458	4200	-5.398	0.000
7	31	87.19	2703				
8	88	44.86	3934	117	3858	-1.707	0.088
11	5	69.80	328				
8	88	424.88	36940	36487	301713	-0.552	0.581
12	736	409.64	301713				
8	88	85.26	7332	3591	7332	-1.728	0.084
13	98	98.86	9688				
7	31	16.30	509	62	77	-0.709	0.478
11	5	16.40	77				
7	31	627.34	19305	4007	275223	-6.125	0.000
12	736	573.64	275223				
7	31	90.26	2798	736	5887	-4.136	0.000
13	98	57.31	5607				
11	5	549.30	2748	848	272185	-1.868	0.062
12	736	589.79	272185				
11	5	98.20	341	164	5015	-1.383	0.214
13	98	51.17	5015				
12	736	486.51	300933	28447	300983	-2.954	0.003
13	98	485.02	47531				

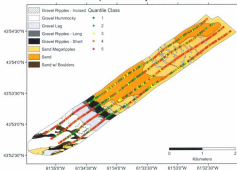
Appendix C: Quantile Classification Maps

Quantile breaks classification of 38 kHz backscatter data
Preferred Study Area



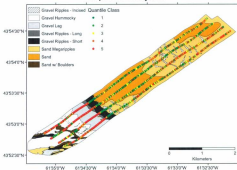
Results of the quantile classification of the 38 kHz backscatter values in the preferred study area overlaid on the interpreted sediment units.

**Quantile breaks classification of 38 kHz backscatter data
Non-preferred Study Area**



Results of the quantile classification of the 38 kHz backscatter values in the non-preferred study area overlaid on the interpreted sediment units.

**Quantile breaks classification of 120 kHz backscatter data
Non-preferred Study Area**

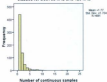


Results of the quantile classification of the 120 kHz backscatter values in the non-preferred study area overlaid on the interpreted sediment units.

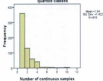
Appendix D: Frequency Differences Based on Classified Backscatter Values

Summary of the differencing of the quantile classification results (Preferred Study Area).

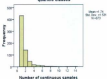
Spatial continuity of samples with equal quantile classes for both 36 kHz and 120 kHz



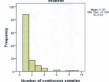
Spatial continuity of samples with higher 36 kHz quantile classes



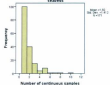
Spatial continuity of samples with higher 120 kHz quantile classes



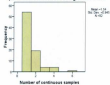
Spatial continuity of samples with equal quantile classes for both 36 kHz and 120 kHz associated over gravel seabeds



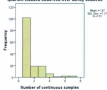
Spatial continuity of samples with equal quantile classes for both 38 kHz and 120 kHz observed over sandy seabeds



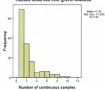
Spatial continuity of samples with higher 38 kHz quantile classes observed over gravel seabeds



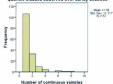
Spatial continuity of samples with higher 38 kHz quantile classes observed over sandy seabeds



Spatial continuity of samples with higher 120 kHz classes observed over gravel seabeds

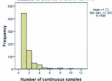


Spatial continuity of samples with higher 120 kHz quantile classes observed over sandy seabeds

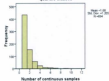


Summary of the differencing of the quantile classification results (Non-preferred Study Area).

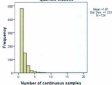
Spatial continuity of samples with equal quantile classes for both 38 kHz and 120 kHz



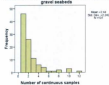
Spatial continuity of samples with higher 38 kHz quantile classes



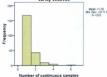
Spatial continuity of samples with higher 120 kHz quantile classes



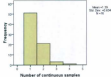
Spatial continuity of samples with equal quantile classes for both 38 kHz and 120 kHz observed over gravel seabeds



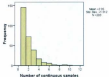
Spatial continuity of samples with equal quantile classes for both 38 kHz and 120 kHz observed over sandy seabeds



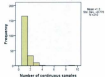
Spatial continuity of samples with higher 38 kHz quantile classes observed over gravel seabeds



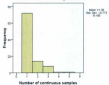
Spatial continuity of samples with higher 30 kHz
quantile classes observed over sandy seabeds



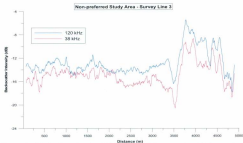
Spatial continuity of samples with higher 120 kHz
classes observed over sandy seabeds



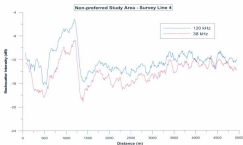
Spatial continuity of samples with higher 120 kHz
classes observed over gravel seabeds



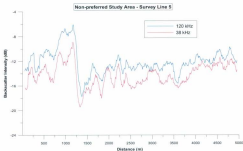
Appendix E: Frequency Differences Based on Continuous Backscatter Values



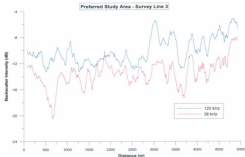
Line graph of backscatter values from both frequencies in the non-preferred study area along survey line 3 (surveyed from northeast to southwest). Data were smoothed using a running average of 9.



Line graph of backscatter values from both frequencies in the non-preferred study area along survey line 4 (surveyed from southwest to northeast). Data were smoothed using a running average of 9.



Line graph of backscatter values from both frequencies in the non-preferred study area along survey line 5 (surveyed from southwest to northeast). Data were smoothed using a running average of 9.



Line graph of backscatter values from both frequencies in the preferred study area along survey line 3 (surveyed from northeast to southwest). Data were smoothed using a running average of 9.

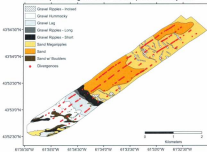


Line graph of backscatter values from both frequencies in the preferred study area along survey line 4 (surveyed from southwest to northeast). Data were smoothed using a running average of 9.

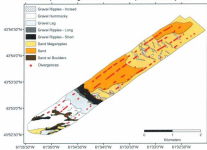


Line graph of backscatter values from both frequencies in the preferred study area along survey line 5 (surveyed from southwest to northeast). Data were smoothed using a running average of 9.

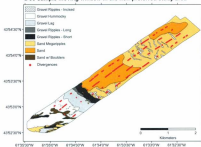
Spatial distribution of frequency divergences based on a 15-sample moving window in the non-preferred study area



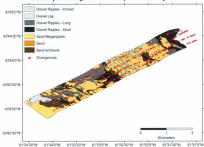
Spatial distribution of frequency divergences based on a 25-sample moving window in the non-preferred study area



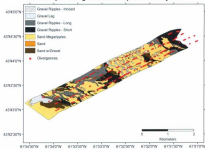
Spatial distribution of frequency divergences based on a 35-sample moving window in the non-preferred study area



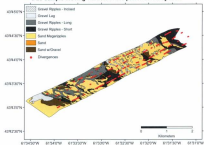
Spatial distribution of frequency divergences based on a 5-sample moving window in the preferred study area

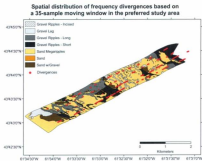


Spatial distribution of frequency divergences based on a 15-sample moving window in the preferred study area



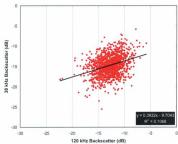
Spatial distribution of frequency divergences based on a 25-sample moving window in the preferred study area



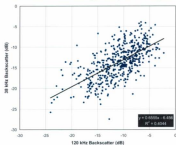


Appendix F: Multivariate Analysis

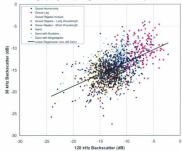
Scatter Plot of 38 kHz and 120 kHz Backscatter of ISUs with Dominant Sand Sediment Type - Non-preferred Study Area



Scatter Plot of 38 kHz and 120 kHz Backscatter of ISUs with Dominant Gravel Sediment Type - Non-preferred Study Area



Scatter Plot of 38 kHz and 120 kHz Frequency Backscatter grouped by Corresponding ISUs - Non-preferred Study Area



[illegible]

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80																				

Preferred Study Areas

[illegible]

[illegible]

[illegible]

[illegible]

